FORM PTO-1390 (REV 5-93)	U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEY DOCKET NO. P108172-00022
	TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371	DATE: December 4, 2000
		U.S. APPLN. NO. (IF KNOWN, SEE 37 C.F.R. 1.5) 09/701395 Not yet assigned

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/US99/12121	June 2, 1999	June 2, 1998

TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF

APPLICANT(S) FOR DO/EO/US: Francis CUNNINGHAM and Zairen SUN

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371. (THE BASIC FILING FEE IS ATTACHED)

2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.

3. ☐ This express request to begin national examination procedures [35 U.S.C. 371(f)] at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).

4. ☒ A proper demand for International Preliminary Amendment was made by the 19th month from the earliest claimed priority date.

5. ☒ A copy of the International Application as filed [35 U.S.C. 371(c)(2)]
   a. ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☒ has been transmitted by the International Bureau.
   c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).

6. ☐ A translation of the International Application into English [35 U.S.C. 371(c)(2)].

7. ☒ Amendments to the claims of the International Application under PCT Article 19 [35 U.S.C. 371(c)(3)]
   a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☒ have been transmitted by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☐ have not been made and will not be made.

8. ☐ A translation of the amendments to the claims under PCT Article 19 [35 U.S.C. 371(c)(3)].

9. ☐ An oath or declaration of the inventor(s) [35 U.S.C. 371(c)(4)].

10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 [35 U.S.C. 371(c)(5)].

Items 11 - 16 below concern other document(s) or information included:

11. An Information Disclosure Statement under 37 C.F.R. 1.97 and 1.98.

12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 C.F.R. 3.28 and 3.31 is included.

13. ☐ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.

14. ☐ A substitute specification.

15. ☐ A change of power of attorney and/or address letter.

16. ☒ Other items or information: Copies of International Prel. Examination Reports (2); PCT/IB/308; PCT/IPEA/408; PCT/IB/332; Copy of Response to Written Opinion dated June 14, 2000; Statement; Computer readable form and paper copy of sequence listing; copy of published application (WO 99/63055)
    ~~CERTIFIED~~.
    Drawings (45 sheets)

09/701395

525 Rec'd PCT/PTO 04 DEC 2000

U.S. APPL'N. NO. (IF KNOWN SEE 37 C.F.R. 1.50) Not yet assigned	INTERNATIONAL APPLICATION NO. PCT/US99/12121	ATTORNEY DOCKET NO. 108172-00022 DATE: December 4, 2000

17. ☒ The following fees are submitted:

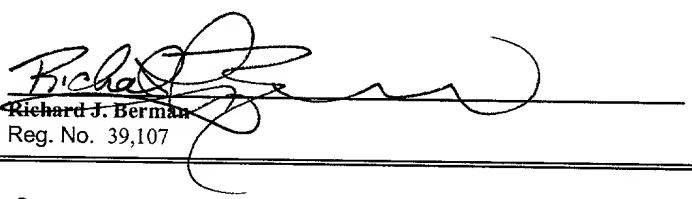
	CALCULATIONS	PTO USE ONLY
Basic National Fee [37 C.F.R. 1.492(a)(1)-(5)]:		
Search Report has been prepared by the EPO or JPO.....$860.00		
International preliminary examination fee paid to USPTO (37 C.F.R. 1.482).....$690.00		
No international preliminary examination fee paid to USPTO (37 C.F.R. 1.482) but international search fee paid to USPTO [37 C.F.R. 1.445(a)(2)].....$710.00		
Neither international preliminary examination fee (37 C.F.R. 1.482) or international search fee [37 C.F.R. 1.445(a)(2)] paid to USPTO.....$1,000.00		
International preliminary examination fee paid to USPTO (37 C.F.R. 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4).....$ 100.00		

ENTER APPROPRIATE BASIC FEE AMOUNT =				$ 690.00	
Surcharge of $130.00 for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date [37 C.F.R. 1.492(e)].				$	
Claims	Number Filed	Number Extra	Rate		
Total Claims	8 - 20 =	00	X $ 18.00	$	
Independent Claims	2 - 3 =	00	X $ 80.00	$	
Multiple dependent claim(s) (if applicable)			+ $270.00	$	
TOTAL OF ABOVE CALCULATIONS =				$ 690.00	
Reduction by one-half for filing by small entity, if applicable. Verified Small Entity statement must also be filed. (Note 37 C.F.R. 1.9, 1.27, 1.28).				$	
SUBTOTAL =				$ 690.00	
Processing fee of $130.00 for furnishing the English translation later the ☐ 20 ☐ 30 months from the earliest claimed priority date [37 C.F.R. 1.492(f)]. +				$	
TOTAL NATIONAL FEE =				$ 690.00	
Fee for recording the enclosed assignment [37 C.F.R. 1.21(h)]. The assignment must be accompanied by an appropriate cover sheet (37 C.F.R. 3.28, 3.31). $40.00 per property +				$	
TOTAL FEES ENCLOSED =				$ 690.00	
				Amount to be refunded	$
				Charged	$

a. ☐ A check in the amount of $ to cover the above fees is enclosed.
b. ☒ Please charge my Deposit Account No. 01-2300 in the amount of $690.00 to cover the above fee.
A duplicate copy of this sheet is enclosed.
c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 01-2300.

NOTE: Where an appropriate time limit under 37 C.F.R. 1.494 or 1.495 has not been met, a petition to revive [37 C.F.R. 1.137(a) or (b)] must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:
Arent Fox Kintner Plotkin & Kahn
1050 Connecticut Avenue, N.W.
Suite 600
Washington, D.C. 20036-5339
Tel: (202) 857-6000 Fax: (202) 638-4810



Richard J. Berman
Reg. No. 39,107

PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of:

CUNNINGHAM, Jr.

Appln. No.: 09/701,395

Filed: December 4, 2000 Attorney Dkt. No.: 108172-00022

For: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF

PRELIMINARY AMENDMENT

Commissioner for Patents
Washington, D.C. 20231 June 18, 2001

Sir:

Prior to [calculation of the filing fees and] initial examination of the application, please amend the above-identified application as follows:

IN THE SPECIFICATION:

Before Line 1, page 1 insert

--CROSS-REFERENCE TO RELATED APPLICATION

This application is a National Stage entry of International Application No. PCT/US99/12121, filed June 2, 1999, the entire specification claims and drawings of which are incorporated herewith by reference. --

REMARKS

In the event that any fees are due with respect to the filing of this paper, please charge our Deposit Account No. 01-2300.

Respectfully submitted,

Richard J. Berman
Registration No. 39,107

ARENT FOX KINTNER PLOTKIN & KAHN, PLLC
Customer No. 004372
1050 Connecticut Avenue, N.W., Suite 600
Washington, D.C. 20036-5339
Tel: (202) 857-6000; Fax: (202) 638-4810
RJB/ccd

GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention describes nucleic acid sequences for eukaryotic genes encoding ϵ lycopene ε-cyclase (also known as ε-cyclase and ε lycopene cyclase), isopentenyl pyrophosphate isomerase (IPP) and β-carotene hydroxylase as well as vectors containing the same and hosts transformed with said vectors. The present invention also provides methods for augmenting the accumulation of carotenoids, changing the composition of the carotenoids, and producing novel and rare carotenoids. The present invention provides methods for controlling the ratio or relative amounts of various carotenoids in a host. The invention also relates to modified lycopene ϵ-cyclase, IPP isomerase and β-carotene hydroxylase. Additionally, the present invention provides a method for screening for genes and cDNAs encoding enzymes of carotenoid biosynthesis and metabolism.

Background of the Invention

Carotenoid pigments with cyclic endgroups are essential components of the photosynthetic apparatus in oxygenic photosynthetic organisms (*e.g.*, cyanobacteria, algae and plants; Goodwin, 1980). The symmetrical bicyclic yellow carotenoid pigment β-carotene (or, in rare cases, the asymmetrical bicyclic α-carotene) is intimately associated with the photosynthetic reaction centers and plays a vital role in protecting against potentially lethal photooxidative damage (Koyama, 1991). β-carotene and other carotenoids derived from it or from α-carotene also serve as light-harvesting pigments (Siefermann-Harms, 1987), are involved in the thermal dissipation of excess light energy captured by the light-harvesting antenna (Demmig-Adams & Adams, 1992), provide substrate for the biosynthesis of the plant growth regulator abscisic acid (Rock & Zeevaart, 1991; Parry & Horgan, 1991), and are precursors of vitamin A in human and animal diets (Krinsky, 1987). Plants also exploit carotenoids as coloring agents in flowers and fruits to attract pollinators and agents of seed dispersal (Goodwin, 1980). The color provided by carotenoids is also of agronomic value in a number of important crops. Carotenoids are currently harvested from a variety of organisms, including plants, algae, yeasts, cyanobacteria and bacteria, for use as pigments in food and feed.

The probable pathway for formation of cyclic carotenoids in plants, algae and cyanobacteria is illustrated in Figure 1. Two types of cyclic endgroups or rings are commonly found in higher plant carotenoids, these are referred to as the β (*beta*) and ε (*epsilon*) rings (Fig. 3). The precursor acyclic endgroup (no ring structure) is referred to as the Ψ (*psi*) endgroup. The β and ε endgroups differ only in the position of the double bond in the ring. Carotenoids with two β rings are ubiquitous, and those with one β and one ε ring are common, but carotenoids with two ε rings are uncommon. β-carotene (Fig. 1) has two β-endgroups and is a symmetrical compound that is the precursor of a number of other important plant carotenoids such as zeaxanthin and violaxanthin (Fig. 2).

Genes encoding enzymes of carotenoid biosynthesis have previously been isolated from a variety of sources including bacteria (Armstrong et al., 1989, Mol. Gen. Genet. 216, 254-268; Misawa et al., 1990, J. Bacteriol., 172, 6704-12), fungi (Schmidhauser et al., 1990, Mol. Cell. Biol. 10, 5064-70), cyanobacteria (Chamovitz et al., 1990, Z. Naturforsch, 45c, 482-86; Cunningham et al., 1994) and higher plants (Bartley et al., Proc. Natl. Acad. Sci USA 88, 6532-36; Martinez-Ferez & Vioque, 1992, Plant Mol. Biol. 18, 981-83). Many of the isolated enzymes show a great diversity in structure, function and inhibitory properties between sources. For example, phytoene desaturases from the cyanobacterium *Synechococcus* and from higher plants and green algae carry out a two-step desaturation to yield ζ-carotene as a reaction product. In plants and cyanobacteria a second enzyme (ζ-carotene desaturase), similar in amino acid sequence to the phytoene desaturase, catalyzes two additional desaturations to yield lycopene. In contrast, a single desaturase enzyme from *Erwinia herbicola* and from other bacteria introduces all four double bonds required to form lycopene. The *Erwinia* and other bacterial desaturases bear little amino acid sequence similarity to the plant and cyanobacterial desaturase enzymes, and are thought to be of unrelated ancestry. Therefore, even with a gene in hand from one source, it may be difficult to identify a gene encoding an enzyme of similar function in another organism. In particular, the sequence similarity between certain of the prokaryotic and eukaryotic genes encoding enzymes of carotenoid biosynthesis is quite low.

Further, the mechanism of gene expression in prokaryotes and eukaryotes appears to differ sufficiently such that one cannot expect that an isolated eukaryotic gene will be properly expressed in a prokaryotic host.

The difficulties in isolating genes encoding enzymes with similar functions is exemplified by recent efforts to isolate the gene encoding the enzyme that catalyzes the formation of β-carotene from the acyclic precursor lycopene. Although a gene encoding an enzyme with this function had been isolated from a bacterium, it had not been isolated from any photosynthetic procaryote or from any eukaryotic organism. The isolation and characterization of the enzyme catalyzing formation of β-carotene in the cyanobacterium *Synechococcus* PCC7942 was described by the present inventors and others (Cunningham et al., 1993 and 1994). The amino acid sequence similarity of the cyanobacterial enzyme to the various bacterial lycopene β–cyclases is so low (*ca.* 18-25% overall; Cunningham et al., 1994) that there is much uncertainty as to whether they share a common ancestry or, instead, represent an example of convergent evolution.

The need remains for the isolation of eukaryotic and prokaryotic genes and cDNAs encoding polypeptides involved in the carotenoid biosynthetic pathway, including those encoding a lycopene ϵ-cyclase, IPP isomerase and β-carotene hydroxylase. There remains a need for methods to enhance the production of carotenoids, to alter the composition of carotenoids, and to reduce or eliminate carotenoid production. There also remains a need in the art for methods for screening for genes and cDNAs encoding enzymes of carotenoid biosynthesis and metabolism.

SUMMARY OF THE INVENTION

Accordingly, a first object of this invention is to provide purified and/or isolated nucleic acids which encode enzymes involved in carotenoid biosynthesis; in particular, lycopene ϵ-cyclase , IPP isomerase and β-carotene hydroxylase.

A second object of this invention is to provide purified and/or isolated nucleic acids which encode enzymes which produce novel or uncommon carotenoids.

A third object of the present invention is to provide vectors containing said genes.

A fourth object of the present invention is to provide hosts transformed with said vectors.

Another object of the present invention is to provide hosts which accumulate novel or uncommon carotenoids or which accumulate greater amounts of specific or total carotenoids.

Another object of the present invention is to provide hosts with inhibited and/or altered carotenoid production.

Another object of this invention is to secure the expression of eukaryotic carotenoid-related genes in a recombinant prokaryotic host.

Yet another object of the present invention is to provide a method for screening for eukaryotic and prokaryotic genes and cDNAs which encode enzymes involved in carotenoid biosynthesis and metabolism.

An additional object of the invention is to provide a method for manipulating carotenoid biosynthesis in photosynthetic organisms by inhibiting the synthesis of certain enzymatic products to cause accumulation of precursor compounds.

Another object of the invention is to provide modified lycopene ϵ-cyclase, IPP isomerase and β-carotene hydroxylase.

These and other objects of the present invention have been realized by the present inventors as described below.

A subject of the present invention is an isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity and having the amino acid sequence of SEQ ID NOS: 2, 4, 14-21 or 23-27.

The invention also includes vectors which comprise any of the nucleic acid sequences listed above, and host cells transformed with such vectors.

Another subject of the present invention is a method of producing or enhancing the production of a carotenoid in a host cell, comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence to produce the protein.

Yet another subject of the present invention is a method of modifying the production of carotenoids in a host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which produces an RNA and/or encodes for a protein which modifies lycopene ϵ-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity, relative to an untransformed host cell, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to the untransformed host cell.

The present invention also includes a method of expressing, in a host cell, a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity, the method comprising inserting into the host cell a vector comprising the heterologous nucleic acid sequence, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence.

Also included is a method of expressing, in a host cell, a heterologous nucleic acid sequence which encodes for a protein which modifies lycopene ϵ-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity in the host cell, relative to an untransformed host cell, the method comprising inserting into the host cell a vector comprising the heterologous nucleic acid sequence, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence.

Another subject of the present invention is a method for screening for genes and cDNAs which encode enzymes involved in carotenoid biosynthesis and metabolism.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete appreciation of the invention and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawings, wherein:

Figure 1 is a schematic representation of the putative pathway of β-carotene biosynthesis in cyanobacteria, algae and plants. The enzymes catalyzing various steps are indicated at the left. Target sites of the bleaching herbicides NFZ and MPTA are also indicated at the left. Abbreviations: DMAPP, dimethylallyl pyrophosphate; FPP, farnesyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate; GPP, geranyl pyrophosphate; IPP, isopentenyl pyrophosphate; LCY, lycopene cyclase; MVA, mevalonic acid; MPTA, 2-(4-methylphenoxy)triethylamine hydrochloride; NFZ, norflurazon; PDS, phytoene desaturase; PSY, phytoene synthase; ZDS, ζ-carotene desaturase; PPPP, prephytoene pyrophosphate.

Figure 2 depicts possible routes of synthesis of cyclic carotenoids and common plant and algal xanthophylls (oxycarotenolds) from neurosporene. Demonstrated activities of the β- and ϵ-cyclase enzymes of *A. thaliana* are indicated by bold arrows labelled with β or ϵ respectively. A bar below the arrow leading to ϵ-carotene indicates that the enzymatic

activity was examined but no product was detected. The steps marked by an arrow with a dotted line have not been specifically examined. Conventional numbering of the carbon atoms is given for neurosporene and α-carotene. Inverted triangles (▼) mark positions of the double bonds introduced as a consequence of the desaturation reactions.

Figure 3 depicts the carotene endgroups which are found in plants.

Figure 4 is a DNA sequence and the predicted amino acid sequence of a lycopene ϵ-cyclase cDNA isolated from *A. thaliana* (SEQ ID NOS: 1 and 2). These sequences were deposited under Genbank accession number U50738. This cDNA is incorporated into the plasmid pATeps.

Figure 5 is a DNA sequence encoding the β-carotene hydroxylase isolated from *A. thaliana* (SEQ ID NO: 3). This cDNA is incorporated into the plasmid pATOHB.

Figure 6 is an alignment of the predicted amino acid sequences of *A. thaliana* β-carotene hydroxylase (SEQ ID NO: 4) with those of the bacterial β-carotene hydroxylase enzymes from *Alicalgenes sp.* (SEQ ID NO: 5) (Genbank D58422), *Erwinia herbicola* Eho10 (SEQ ID NO.: 6) (GenBank M872280), *Erwinia uredovora* (SEQ ID NO.: 7) (GenBank D90087) and *Agrobacterium aurianticum* (SEQ ID NO.: 8) (GenBank D58420). A consensus sequence is also shown. All five genes are identical where a capital letter appears in the consensus. A lowercase letter indicates that three of five, including *A. thaliana*, have the identical residue. TM; transmembrane.

Figure 7 is a DNA sequence of a cDNA encoding an IPP isomerase isolated from *A. thaliana* (SEQ ID NO: 9). This cDNA is incorporated into the plasmid pATDP5.

Figure 8 is a DNA sequence of a second cDNA encoding another IPP isomerase isolated from *A. thaliana* (SEQ ID NO: 10). This cDNA is incorporated into the plasmid pATDP7.

Figure 9 is a DNA sequence of a cDNA encoding an IPP isomerase isolated from *Haematococcus pluvialis* (SEQ ID NO: 11). This cDNA is incorporated into the plasmid pHP04.

Figure 10 is a DNA sequence of a second cDNA encoding another IPP isomerase isolated from *Haematococcus pluvialis* (SEQ ID NO: 12). This cDNA is incorporated into the plasmid pHP05.

Figure 11 is an alignment of the amino acid sequences predicted by IPP isomerase cDNAs isolated from *A. thaliana (SEQ ID NO.: 16 and 18), H. pluvialis (SEQ ID NOS.: 14*

and 15), Clarkia breweri (SEQ ID NO.: 17) (See, Blanc & Pichersky, Plant Physiol. (1995) 108:855; Genbank accession no. X82627) and *Saccharomyces cerevisiae* (SEQ ID NO.: 19) (Genbank accession no. J05090).

Figure 12 is a DNA sequence of the cDNA encoding an IPP isomerase isolated from *Tagetes erecta* (marigold; SEQ ID NO: 13). This cDNA is incorporated into the plasmid pPMDP1. xxx's denote a region not originally sequenced. Figure 21A shows the complete marigold sequence.

Figure 13 is an alignment of the consensus sequence of four plant β-cyclases (SEQ ID NO.: 20) with the *A. thaliana* lycopene ϵ-cyclase (SEQ ID NO.: 21). A capital letter in the plant β consensus is used where all four β-cyclase genes predict the same amino acid residue in this position. A small letter indicates that an identical residue was found in three of the four. Dashes indicate that the amino acid residue was not conserved and dots in the sequence denote a gap. A consensus for the aligned sequences is given, in capital letters below the alignment, where the β- and ϵ-cyclases have the same amino acid residue. Arrows indicate some of the conserved amino acids that will be used as junction sites for construction of chimeric cyclases with novel enzymatic activities. Several regions of interest including a sequence signature indicative of a dinucleotide-binding motif and two predicted transmembrane (TM) helical regions are indicated below the alignment and are underlined.

Figure 14 shows the nucleotide (SEQ ID NO:22) and amino acid sequences (SEQ ID NO:23) of the *Adonis palaestina* (pheasant's eye) ϵ-cyclase cDNA #5.

Figure 15A shows the nucleotide (SEQ ID NO:24) and amino acid sequences (SEQ ID NO:25) of a potato ϵ-cyclase cDNA. Figure 15B shows the amino acid sequence (SEQ ID NO:26) of a chimeric lettuce/potato lycopene ϵ-cyclase. Amino acids in lower case are from the lettuce cDNA and those in upper case are from the potato cDNA. The product of this chimeric cDNA has e-cyclase activity and converts lycopene to the monocyclic δ-carotene.

Figure 16 shows a comparison between the amino acid sequences of the *Arabidopsis* ϵ-cyclase (SEQ ID NO:27) and the potato ϵ-cyclase (SEQ ID NO:25).

Figure 17A shows the nucleotide sequence of the *Adonis palaestina* Ipi1 (SEQ ID NO:28) and Figure 17B shows the nucleotide sequence of the *Adonis palaestina* Ipi2 (SEQ ID NO: 29).

Figure 18A shows the nucleotide sequence of the *Haematoccus pluvialis* Ipi1 (SEQ ID NO:11) and Figure 18B shows the nucleotide sequence of the *Haematoccus pluvialis* Ipi2 (SEQ ID NO:30).

Figure 19A shows the nucleotide sequence of the *Lactuca sativa (romaine lettuce)* Ipi1 (SEQ ID NO:31) and Figure 19B shows the nucleotide sequence of the *Lactuca sativa* Ipi2 (SEQ ID NO: 32).

Figure 20 shows the nucleotide sequence of the *Chlamydomonas reinhardtii* Ipi1 (SEQ ID NO:33).

Figure 21A shows the nucleotide sequence of the *Tagetes erecta* (marigold) Ipi1 (SEQ ID NO:34) and Figure 21B shows the nucleotide sequence of the *Oryza sativa* (rice) Ipi1 (SEQ ID NO:35).

Figure 22 shows a amino acid sequence alignment of various plant and green algal isopentenyl isomerases (IPI) (SEQ ID NOS:16, 36-45).

Figure 23 shows a comparison between *Adonis palaestina* ϵ-cyclase cDNA #3 and cDNA #5 nucleotide sequences.

Figure 24 shows a comparison between *Adonis palaestina* ϵ-cyclase cDNA #3 and cDNA #5 predicted amino acid sequences.

Figure 25 shows a sequence alignment of various plant β- and ϵ-cyclases. Those sequences outlined in grey denote identical sequences among the ϵ-cyclases. Those sequences outlined in black denote identical sequences among both the β- and ϵ-cyclases.

Figure 26 shows a sequence alignment of the plant ϵ-cyclases from Figure 25. Those sequences outlined in black denote identical sequences among the ϵ-cyclases.

Figure 27 is a dendrogram or "tree" illustrating the degree of amino acid sequence similarity for various lycopene β- and ϵ-cyclases.

Figure 28 shows a comparison between Arabidopsis ϵ-cyclase and lettuce ϵ-cyclase predicted amino acid sequences.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention includes an isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity and having the amino acid sequence of SEQ ID NOS: 2, 4, 14-21, 23 or 25-27. Nucleic acids encoding lycopene ϵ-cyclase, β-carotene hydroxylase and IPP

isomerases have been isolated from several genetically distant sources.

The present inventors have isolated nucleic acids encoding the enzyme IPP isomerase, which catalyzes the reversible conversion of isopentenyl pyrophosphate (IPP) to dimethylallyl pyrophosphate (DMAPP). IPP isomerase cDNAs were isolated from the plants *A. thaliana, Tagetes erecta* (marigold), *Adonis palaestina* (pheasant's eye), *Lactuca sativa* (romaine lettuce) and from the green algae *H. pluvialis* and *Chlamydomonas reinhardtii.* Alignments of the amino acid sequences predicted by some of these cDNAs are shown in Figures 12 and 22. Plasmids containing some of these cDNAs were deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession numbers 98000 (pHP05 - *H. pluvialis*); 98001 (pMDP1 - marigold); 98002 (pATDP7 - *A. thaliana*) and 98004 (pHP04 - *H. pluvialis*).

The present inventors have also isolated nucleic acids encoding the enzyme β-carotene hydroxylase, which is responsible for hydroxylating the β-endgroup in carotenoids. The nucleic acid of the present invention is shown in SEQ ID NO: 3 and Figure 5. The full length cDNA product hydroxylates both end groups of β-carotene as do products of cDNAs which encode proteins truncated by up to 50 amino acids from the N-terminus. Products of genes which encode proteins truncated between about 60-110 amino acids from the N-terminus preferentially hydroxylate only one ring. A plasmid containing this gene was deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession number 98003 (pATOHB - *A. thaliana*).

The present inventors have also isolated nucleic acids encoding the enzyme lycopene ε-cyclase, which is responsible for the formation of ε-endgroups in carotenoids. The *A. thaliane* ε-cyclase adds an ε ring to only one end of the symmetrical lycopene while the related β-cyclase adds a ring at both ends. The *A. thaliana* cDNA of the present invention is shown in Figure 4 and SEQ ID NO: 1. A plasmid containing this gene was deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville MD 20852 on March 4, 1996 under ATCC accession number 98005 (pATeps - *A. thaliana*).

In addition, lycopene ε-cyclases have been identified in lettuce and in *Adonis palaestina* (cDNA #5) which encode enzymes that convert lycopene to the bicyclic ε-carotene (ε,ε-carotene). An additional cDNA from *Adonis palaestina* (cDNA #3) encodes a lycopene ε-cyclase which converts lycopene into δ-carotene (ε,ψ-carotene) and differs from the lycopene ε-cyclase which forms bicyclic ε-carotene (ε,ε-carotene) by only 5 amino acids.

One or more of these amino acids may be modified by alteration of the nucleotide sequence in the #5 cDNA to obtain an enzyme which forms the bicyclic ε,ε-carotene. The sequences of the *Adonis palaestina* and *Arabidopsis thaliana* ε-cyclases have about 70% nucleotide identity and about 72% amino acid identity.

Initial experiments by the inventors with chimeric genes indicated that the part of the ε-cyclase which is responsible for adding 2 ε rings to form ε,ε-carotene is the carboxy terminal portion of the gene. The lettuce ε-cyclase adds two ε rings to form ε,ε-carotene. A DNA encoding a partial potato ε-cyclase (missing its amino terminal portion), when combined with an amino terminal region from the lettuce ε-cyclase gene, produces a monocyclic δ-carotene (ε,ψ-carotene). With the discovery of the differences between the *Adonis palaestina* clone #3 and clone #5, the specific amino acids responsible for the addition of an extra ε ring have been identified (Figure 24). Specifically, amino acid 55 is Thr in clone #3 and Ser in clone #5, amino acid 210 is Asn in clone #3 and Asp in clone #5, amino acid 231 is Asp in clone #3 and Glu in clone #5, amino acid 352 is Ile in clone #3 and Val in clone #5, and amino acid 524 is Lys in clone #3 and Arg in clone #5. It can be appreciated that these changes are quite conservative, as only one change, at amino acid 210, changes the charge of the protein.

Thus, it is clear that the nucleic acids of the invention encoding the enzymes as presently disclosed may be altered to increase a particularly desirable property of the enzyme, to change a property of the enzyme, or to diminish an undesirable property of the enzyme. Such modifications can be by deletion, substitution, or insertion of one or more amino acids, and can be performed by routine enzymatic manipulation of the nucleic acid encoding the enzyme (such as by restriction enzyme digestion, removal of nucleotides by mung bean nuclease or *Bal*31, insertion of nucleotides by Klenow fragment, and by religation of the ends), by site-directed mutagenesis, or may be accidental, such as by low fidelity PCR or those obtained through mutations in hosts that are producers of the enzymes. These techniques as well as other suitable techniques are well known in the art.

Mutations can be made in the nucleic acids of the invention such that a particular codon is changed to a codon which codes for a different amino acid. Such a mutation is generally made by making the fewest nucleotide changes possible. A substitution mutation of this sort can be made to change an amino acid in the resulting protein in a non-conservative manner (i.e., by changing the codon from an amino acid belonging to a grouping

of amino acids having a particular size or characteristic to an amino acid belonging to another grouping) or in a conservative manner (i.e., by changing the codon from an amino acid belonging to a grouping of amino acids having a particular size or characteristic to an amino acid belonging to the same grouping). Such a conservative change generally leads to less change in the structure and function of the resulting protein. A non-conservative change is more likely to alter the structure, activity or function of the resulting protein. The present invention should be considered to include sequences containing conservative changes which do not significantly alter the activity or binding characteristics of the resulting protein.

The following is one example of various groupings of amino acids:

Amino acids with nonpolar R groups: Alanine, Valine, Leucine, Isoleucine, Proline, Phenylalanine, Tryptophan and Methionine.

Amino acids with uncharged polar R groups: Glycine, Serine, Threonine, Cysteine, Tyrosine, Asparagine and Glutamine.

Amino acids with charged polar R groups (negatively charged at Ph 6.0): Aspartic acid and Glutamic acid.

Basic amino acids (positively charged at pH 6.0): Lysine, Arginine and Histidine.

Another grouping may be those amino acids with phenyl groups: Phenylalanine, Tryptophan and Tyrosine.

Another grouping may be according to molecular weight (i.e., size of R groups).

Particularly preferred substitutions are:

- Lys for Arg and vice versa such that a positive charge may be maintained;
- Glu for Asp and vice versa such that a negative charge may be maintained;
- Ser for Thr such that a free -OH can be maintained; and
- Gln for Asn such that a free NH_2 can be maintained.

Amino acid substitutions may also be introduced to substitute an amino acid with a particularly preferable property. For example, a Cys may be introduced to provide a potential site for disulfide bridges with another Cys. A His may be introduced as a particularly "catalytic" site (i.e., His can act as an acid or base and is the most common amino acid in biochemical catalysis). Pro may be introduced because of its particularly planar structure, which induces β-turns in the protein's structure.

It is clear that certain modifications of SEQ ID NOS: 2, 4, 14-21, 23 or 25-27 can take place without destroying the activity of the enzyme. It is noted especially that truncated

versions of the nucleic acids of the invention are functional. For example, several amino acids (from 1 to about 120) can be deleted from the N-terminus of the lycopene ϵ-cyclases of the invention, and a functional protein can still be produced. This fact is made especially clear from Figure 25, which shows a sequence alignment of several plant ϵ-cyclases. As can be seen from Figure 25, there is an enormous amount of sequence disparity between amino acid sequences 2 to about 50-70 (depending on the particular sequence, since gaps are present). There is less, but also a substantial amount of, sequence dissimilarity between about 50-70 to about 90-120 (depending on the particular sequence). Thereafter, the sequences are fairly conserved, except for small pockets of dissimilarity between about 275-295 to about 285-305 (depending on the particular sequence), and between about 395-415 to about 410-430 (depending on the particular sequence).

The present inventors have found that the amount of the 5' region present in the nucleic acids of the invention can alter the activity of the enzyme. Instead of diminishing activity, truncating the 5' region of the nucleic acids of the invention may result in an enzyme with a different specificity. Thus, the present invention relates to nucleic acids and enzymes encoded thereby which are truncated to within 0-50, preferably 0-25, codons of the 5' initiation codon of their prokaryotic counterparts as determined by alignment maps as discussed below.

For example, when the cDNA encoding *A. thaliana* β-carotene hydroxylase was truncated, the resulting enzyme catalyzed the formation of β-cryptoxanthin as the major product and zeaxanthin as minor product; in contrast to its normal production of zeaxanthin.

The present invention is intended to include those nucleic acid and amino acid sequences in which substitutions, deletions, additions or other modifications have taken place, as compared to SEQ ID NOS: 2, 4, 14-21, 23 or 25-27, without destroying the activity of the enzyme. Preferably, the substitutions, deletions, additions or other modifications take place at the 5' end, or any other of those positions which already show dissimilarity between any of the presently disclosed amino acid sequences (see also Figure 25) or other amino acid sequences which are known in the art and which encode the same enzyme (i.e., lycopene ϵ-cyclase, IPP isomerase or β-carotene hydroxylase).

In each case, nucleic acid and amino acid sequence similarity and identity is measured using sequence analysis software, for example, the Sequence Analysis, Gap, or BestFit software packages of the Genetics Computer Group (University of Wisconsin Biotechnology

Center, 1710 University Avenue, Madison, Wisconsin 53705), MEGAlign (DNAStar, Inc., 1228 S. Park St., Madison, Wisconsin 53715), or MacVector (Oxford Molecular Group, 2105 S. Bascom Avenue, Suite 200, Campbell, California 95008). Such software uses algorithms to match similar sequences by assigning degrees of identity to various substitutions, deletions, and other modifications, and includes detailed instructions as to useful parameters, etc., such that those of routine skill in the art can easily compare sequence similarities and identities. An example of a useful algorithm in this regard is the algorithm of Needleman and Wunsch, which is used in the Gap program discussed above. This program finds the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. Another useful algorithm is the algorithm of Smith and Waterman, which is used in the BestFit program discussed above. This program creates an optimal alignment of the best segment of similarity between two sequences. Optimal alignments are found by inserting gaps to maximize the number of matches using the local homology algorithm of Smith and Waterman.

Conservative (i.e. similar) substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine and leucine; aspartic acid, glutamic acid, asparagine and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine. Substitutions may also be made on the basis of conserved hydrophobicity or hydrophilicity (see Kyte and Doolittle, *J. Mol. Biol.* **157**: 105-132 (1982)), or on the basis of the ability to assume similar polypeptide secondary structure (see Chou and Fasman, *Adv. Enzymol.* **47**: 45-148 (1978)).

If comparison is made between nucleotide sequences, preferably the length of comparison sequences is at least 50 nucleotides, more preferably at least 60 nucleotides, at least 75 nucleotides or at least 100 nucleotides. It is most preferred if comparison is made between the nucleic acid sequences encoding the enzyme coding regions necessary for enzyme activity. If comparison is made between amino acid sequences, preferably the length of comparison is at least 20 amino acids, more preferably at least 30 amino acids, at least 40 amino acids or at least 50 amino acids. It is most preferred if comparison is made between the amino acid sequences in the enzyme coding regions necessary for enzyme activity.

It should be appreciated that also within the scope of the present invention are nucleic acid sequences encoding lycopene ϵ-cyclases, IPP isomerases and β-carotene hydroxylases

which code for enzymes having the same amino acid sequence as SEQ ID NOS: 2, 4, 14-21, 23 or 25-27, but which are degenerate to the nucleic acids specifically disclosed herein.

The amino acid residues described herein are preferred to be in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property of immunoglobulin-binding is retained by the polypeptide.

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook et al, "Molecular Cloning: A Laboratory Manual" (1989); "Current Protocols in Molecular Biology" Volumes I-III [Ausubel, R. M., ed. (1994)]; "Cell Biology: A Laboratory Handbook" Volumes I-III [J. E. Celis, ed. (1994))]; "Current Protocols in Immunology" Volumes I-III [Coligan, J. E., ed. (1994)]; "Oligonucleotide Synthesis" (M.J. Gait ed. 1984); "Nucleic Acid Hybridization" [B.D. Hames & S.J. Higgins eds. (1985)]; "Transcription And Translation" [B.D. Hames & S.J. Higgins, eds. (1984)]; "Animal Cell Culture" [R.I. Freshney, ed. (1986)]; "Immobilized Cells And Enzymes" [IRL Press, (1986)]; B. Perbal, "A Practical Guide To Molecular Cloning" (1984).

The present invention also includes vectors. Suitable vectors according to the present invention comprise a nucleic acid of the invention encoding an enzyme involved in carotenoid biosynthesis or metabolism and a suitable promoter for the host, and can be constructed using techniques well known in the art (for example Sambrook et al., Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing and Wiley Interscience, New York, 1991). Suitable vectors for eukaryotic expression in plants are described in Frey et al., Plant J. (1995) 8(5):693 and Misawa et al, 1994a; incorporated herein by reference. Suitable vectors for prokaryotic expression include pACYC184, pUC119, and pBR322 (available from New England BioLabs, Bevery, MA) and pTrcHis (Invitrogen) and pET28 (Novagen) and derivatives thereof. The vectors of the present invention can additionally contain regulatory elements such as promoters, repressors, selectable markers such as antibiotic resistance genes, etc.

The nucleic acids encoding the carotenoid enzymes as described above, when cloned into a suitable expression vector, can be used to overexpress these enzymes in a plant

expression system or to inhibit the expression of these enzymes. For example, a vector containing the gene encoding lycopene ϵ-cyclase can be used to increase the amount of α-carotene and carotenoids derived from α-carotene (such as lutein and α-cryptoxanthin) in an organism and thereby alter the nutritional value, pharmacology and visual appearance value of the organism.

Therefore, the present invention includes a method of producing or enhancing the production of a carotenoid in a host cell, relative to an untransformed host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which encodes for a protein having lycopene ϵ-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence to produce the protein.

The present invention also includes a method of modifying the production of carotenoids in a host cell, the method comprising inserting into the host cell a vector comprising a heterologous nucleic acid sequence which produces an RNA and/or encodes for a protein which modifies lycopene ϵ-cyclase, IPP isomerase or β-carotene hydroxylase enzyme activity, relative to an untransformed host cell, wherein the heterologous nucleic acid sequence is operably linked to a promoter; and expressing the heterologous nucleic acid sequence in the host cell to modify the production of the carotenoids in the host cell, relative to the untransformed host cell.

The term "modifying the production" means that the amount of carotenoids produced in the host cell can be enhanced, reduced, or left the same, as compared to the untransformed host cell. In accordance with one embodiment of the present invention, the make-up of the carotenoids (i.e., the specific carotenoids produced) is changed *vis a vis* each other, and this change in make-up may result in either a net gain, net loss, or no net change in the total amount of carotenoids produced in the cell. In accordance with another embodiment of the present invention, the production or the biochemical activity of the carotenoids (or the enzymes which catalyze their formation) is enhanced by the insertion of an enzyme-encoding nucleic acid of the invention. In yet another embodiment of the invention, the production or the biochemical activity of the carotenoids (or the enzymes which catalyze their formation) may be reduced or inhibited by a number of different approaches available to those skilled in the art, including but not limited to such methodologies or approaches as anti-sense (e.g.,

Gray et al (1992) Plant Mol. Biol. 19:69-87), ribozymes (e.g., Wegener et al (1994) Mol. Gen. Genet. 245:465-470), co-suppression (e.g., Fray and Grierson (1993) Plant Mol. Biol. 22:589-602), targeted disruption of the gene (e.g., Schaefer et al. (1997) Plant J. 11:1195-1206), intracellular antibodies (e.g., Rondon and Marasco (1997) Ann. Rev. Microbiol. 51:257-283) or whatever other approaches rely on the knowledge or availability of the nucleic acid or amino acid sequences of the invention and/or portions thereof, to thereby reduce accumulation of carotenoids with ϵ rings and compounds derived from them (for ϵ-cyclase inhibition), or carotenoids with hydroxylated β rings and compounds derived from them (for β-hydroxylase inhibition), or, in the case if IPP isomerase, accumulation of any isoprenoid compound.

Preferably, at least a portion of the nucleic acid sequences used in the methods, vectors and host cells of the invention codes for an enzyme having an amino acid sequence which is at least 85% identical, preferably at least 90%, at least 95% or completely identical to SEQ ID NOS: 2, 4, 14-21, 23 or 25-27. Sequence identity is determined as noted above. Preferably, sequence additions, deletions or other modifications are made as indicated above, so as to not affect the function of the particular enzyme.

In a preferred embodiment, vectors are manufactured which contain a DNA encoding a eukaryotic IPP isomerase upstream of a DNA encoding a second eukaryotic carotenoid enzyme. The inventors have discovered that inclusion of an IPP isomerase gene increases the supply of substrate for the carotenoid pathway; thereby enhancing the production of carotenoid endproducts, as compared to a host cell which is not transformed with such a vector. This is apparent from the much deeper pigmentation in carotenoid-accumulating colonies of *E. coli* which also contain one of the aforementioned IPP isomerase genes when compared to colonies that lack this additional IPP isomerase gene. Similarly, a vector comprising an IPP isomerase gene can be used to enhance production of any secondary metabolite of dimethylallyl pyrophosphate and/or isopentenyl pyrophosphate (such as isoprenoids, steroids, carotenoids, etc.). The term "isoprenoid" is intended to mean any member of the class of naturally occurring compounds whose carbon skeletons are composed, in part or entirely, of isopentyl C_5 units. Preferably, the carbon skeleton is of an essential oil, a fragrance, a rubber, a carotenoid, or a therapeutic compound, such as paclitaxel.

A vector containing the cDNA encoding a lycopene ϵ-cyclase of the invention, preferably the lettuce lycopene ϵ-cyclase or Adonis ϵ-cyclase #5, can be used to increase the

amount of bicyclic ε-carotene in an organism and thereby alter the nutritional value, pharmacology and visual appearance value of the organism. In addition, the transformed organism can be used in the formulation of therapeutic agents, for example in the treatment of cancer (see Mayne et al (1996) FASEB J. 10:690-701; Tsushima et al (1995) Biol. Pharm. Bull. 18:227-233).

An antisense strand of a nucleic acid of the invention can be inserted into a vector. For example, the lycopene ε-cyclase gene can be inserted into a vector and incorporated into the genomic DNA of a host, thereby inhibiting the synthesis of ε,β-carotenoids (lutein and α-carotene) and enhancing the synthesis of β,β-carotenoids (zeaxanthin and β-carotene).

The present invention also relates to novel enzymes which are encoded by the amino acid sequences of the invention, or portions thereof.

The present invention also relates to novel enzymes which can transform known carotenoids into novel or uncommon products. Currently ε-carotene (see Figure 2) and γ-carotene are commonly produced only in minor amounts. As described below, an enzyme can be produced which transforms lycopene to γ-carotene and lycopene to ε-carotene. With these products in hand, bulk synthesis of other carotenoids derived from them are possible. For example, ε-carotene can be hydroxylated to form lactucaxanthin, an isomer of lutein (one ε and one β ring) and zeaxanthin (two β rings) where both endgroups are, instead, ε rings.

In addition to novel enzymes produced by truncating the 5' region of known enzymes, as discussed above, novel enzymes which can participate in the formation of unusual carotenoids can be formed by replacing portions of one gene with an analogous sequence from a structurally related gene. For example, β-cyclase and ε-cyclase are structurally related (see Figure 13). By replacing a portion of β-lycopene cyclase with the analogous portion of ε-cyclase, an enzyme which produces γ-carotene will be produced (one β endgroup). Further, by replacing a portion of the lycopene ε-cyclase with the analogous portion of β-cyclase, an enzyme which produces ε-carotene will be produced (with some exceptions, such as the lettuce ε-cyclase, plant ε-cyclases normally produce a compound with one ε-endgroup, δ-carotene). Similarly, β-hydroxylase could be modified to produce enzymes of novel function by creation of hybrids with ε-hydroxylase.

Host systems according to the present invention can comprise any organism that already produces carotenoids or which has been genetically modified to produce carotenoids.

The IPP isomerase genes are more broadly applicable for enhancing production of any product dependent on DMAPP and/or IPP as a precursor.

Organisms which already produce carotenoids include plants, algae, some yeasts, fungi and cyanobacteria and other photosynthetic bacteria. Transformation of these hosts with vectors according to the present invention can be done using standard techniques such as those described in Misawa et al., (1990) supra; Hundle et al., (1993) supra; Hundle et al., (1991) supra; Misawa et al., (1991) supra; Sandmann et al., supra; and Schnurr et al., supra.

Transgenic organisms can be constructed which include the nucleic acid sequences of the present invention (Bird et al, 1991; Bramley et al, 1992; Misawa et al, 1994a; Misawa et al, 1994b; Cunningham et al, 1993). The incorporation of these sequences can allow the controlling of carotenoid biosynthesis, content, or composition in the host cell. These transgenic systems can be constructed to incorporate sequences which allow for the overexpression of the nucleic acids of the present invention. Transgenic systems can also be constructed containing antisense expression of the nucleic acid sequences of the present invention. Such antisense expression would result in the accumulation of the substrates of the substrates of the enzyme encoded by the sense strand.

A method for screening for eukaryotic genes which encode enzymes involved in carotenoid biosynthesis comprises transforming a prokaryotic host with a nucleic acid which may contain a eukaryotic or prokaryotic carotenoid biosynthetic gene; culturing said transformed host to obtain colonies; and screening for colonies exhibiting a different color than colonies of the untransformed host.

Suitable hosts include *E. coli,* cyanobacteria such as *Synechococcus* and *Synechocystis,* alga and plant cells. *E. coli* are preferred.

In a preferred embodiment, the above "color complementation" screening protocol can be enhanced by using mutants which are either (1) deficient in at least one carotenoid biosynthetic gene or (2) overexpress at least one carotenoid biosynthetic gene. In either case, such mutants will accumulate carotenoid precursors.

Prokaryotic and eukaryotic DNA or cDNA libraries can be screened in total for the presence of genes of carotenoid biosynthesis, metabolism and degradation. Preferred organisms to be screened include photosynthetic organisms.

E. coli can be transformed with these eukaryotic cDNA libraries using conventional methods such as those described in Sambrook et al, 1989 and according to protocols described by the vendors of the cloning vectors.

For example, the cDNA libraries in bacteriophage vectors such as lambdaZAP (Stratagene) or lambda ZIPLOX (Gibco BRL) can be excised en masse and used to transform *E.coli*.

Transformed *E. coli* can be cultured using conventional techniques. The culture broth preferably contains antibiotics to select and maintain plasmids. Suitable antibiotics include penicillin, ampicillin, chloramphenicol, etc. Culturing is typically conducted at 15-40°C, preferably at room temperature or slightly above (18-28°C), for 12 hours to 7 days.

Cultures are plated and the plates are screened visually for colonies with a different color than the colonies of the host *E. coli* transformed with the empty plasmid cloning vector. For example, *E. coli* transformed with the plasmid, pAC-BETA (described below), produce yellow colonies that accumulate β-carotene. After transformation with a cDNA library, colonies which contain a different hue than those formed by *E. coli*/pAC-BETA would be expected to contain enzymes which modify the structure or accumulation of β-carotene. Similar E. coli strains can be engineered which accumulate earlier products in carotenoid biosynthesis, such as lycopene, γ-carotene, etc.

Having generally described this invention, a further understanding can be obtained by reference to certain specific examples which are provided herein for purposes of illustration only and are not intended to be limiting unless otherwise specified.

EXAMPLE

I. Isolation of β-carotene hydroxylase

Plasmid Construction

An 8.6kb BglII fragment containing the carotenoid biosynthetic genes of *Erwinia herbicola* was first cloned in the BamHI site of plasmid vector pACYC184 (chloramphenicol resistant), and then a 1.1kb BamHI fragment containing the *E. herbicola* β-carotene hydroxylase (*CrtZ*) was deleted. *E.coli* strains containing the resulting plasmid, pAC-BETA, accumulate ß-carotene and form yellow colonies (Cunningham et al., 1994).

A full length cDNA encoding IPP isomerase of *Haematococcus pluvialis* (HP04) was first excised with *BamH*I and *Kpn*I from pBluescript SK-, and then ligated into the

corresponding sites of the pTrcHisA vector with high-level expression from the *trc* promoter (Invitrogen, Inc.). A fragment containing the IPP isomerase and *trc* promoter was subsequently excised with *EcoR*V and *Kpn*I, treated with the Klenow fragment of DNA polymerase to produce blunt ends, and ligated in the Klenow-treated *Hind*III site of pAC-BETA. *E.coli* cells transformed with this new plasmid pAC-BETA-04 form orange colonies on LB plates (*vs.* yellow for those containing pAC-BETA) and cultures accumulate substantially more β-carotene (*ca.* two fold) than those that contain pAC-BETA.

Screening of an Arabidopsis cDNA Library

Several λ cDNA expression libraries of *Arabidopsis* were obtained from the *Arabidopsis* Biological Resource Center (Ohio State University, Columbus, OH) (Kieber et al., 1993). The λ cDNA libraries were excised *in vivo* using Stratagene's ExAssist SOLR system to produce a phagemid cDNA library wherein each phagemid contained also a gene conferring resistance to the antibiotic ampicillin.

E.coli strain DH10BZIP was chosen as the host cell for the screening and pigment production, although we have also used TOP10F' and XL1-Blue for this purpose. DH10B cells were transformed with plasmid pAC-BETA-04 and were plated on LB agar plates containing chloramphenicol at 50 µg/ml (from United States Biochemical Corporation). The phagemid *Arabidopsis* cDNA library was then introduced into DH10B cells already containing pAC-BETA-04. Transformed cells containing both pAC-BETA-04 and *Arabidopsis* cDNA library phagemids were selected on chloramphenicol plus ampicillin (150 µg/ml) agar plates. Maximum color development occurred after 3 to 7 days incubation at room temperature, and the rare bright yellow colonies were selected from a background of many thousands of orange colonies on each agar plate. Selected colonies were inoculated into 3 ml liquid LB medium containing ampicillin and chloramphenicol, and cultures were incubated at room temperature for 1-2 days, with shaking. Cells were then harvested by centrifugation and extracted with acetone in microfuge tubes. After centrifugation, the pigmented extract was spotted onto silica gel thin-layer chromatography (TLC) plates, and developed with a hexane:ether (1:1, by volume) mobile phases. ß-carotene hydroxylase-encoding cDNAs were identified based on the appearance of a yellow pigment that co-migrated with zeaxanthin on the TLC plates.

Subcloning and Sequencing

The plasmid containing the ß-carotene hydroxylase cDNA was recovered and analyzed by standard procedures (Sambrook et al., 1989). The *Arabidopsis* ß-carotene hydroxylase was sequenced completely on both strands on an automatic sequencer (Applied Biosystems, Model 373A, Version 2.0.1S). The cDNA insert of 0.95kb also was excised and ligated into the a pTrcHis vector. A *Bgl*II restriction site within the cDNA was used to remove that portion of the cDNA that encodes the predicted polypeptide N terminal sequence region that is not also found in bacterial β-carotene hydroxylases (Figure 6). A BglII-XhoI fragment was directionally cloned in BamHI-XhoI digested TrcHis vectors.

Pigment Analysis

A single colony was used to inoculate 50 ml of LB containing ampicillin and chloramphenicol in a 250-ml flask. Cultures were incubated at 28°C for 36 hours with gentle shaking, and then harvested at 5000 rpm in an SS-34 rotor. The cells were washed once with distilled H_2O and resuspended with 0.5 ml of water. The extraction procedures and HPLC were essentially as described previously (Cunningham et al, 1994).

II. Isolation and biochemical analysis of an Arabidopsis lycopene ϵ-cyclase

Plasmid Construction

Construction of plasmids pAC-LYC, pAC-NEUR, and pAC-ZETA is described in Cunningham et al., (1994). In brief, the appropriate carotenoid biosynthetic genes from *Erwinia herbicola*, *Rhodobacter capsulatus*, and *Synechococcus* sp. strain PCC7942 were cloned in the plasmid vector pACYC184 (New England BioLabs, Beverly, MA). Cultures of *E. coli* containing the plasmids pAC-ZETA, pAC-NEUR, and pAC-LYC, accumulate ζ-carotene, neurosporene, and lycopene, respectively. The plasmid pAC-ZETA was constructed as follows: an 8.6-kb BglII fragment containing the carotenoid biosynthetic genes of *E. herbicola* (GenBank M87280; Hundle et al., 1991) was obtained after partial digestion of plasmid pPL376 (Perry et al., 1986; Tuveson et al., 1986) and cloned in the BamHI site of pACYC184 to give the plasmid pAC-EHER. Deletion of adjacent 0.8- and 1.1-kb BamHI-BamHI fragments (deletion Z in Cunningham et al., 1994), and of a 1.1 kB SalI-SalI fragment (deletion X) served to remove most of the coding regions for the *E. herbicola* β-carotene hydroxylase (crtZ gene) and zeaxanthin glucosyltransferase (crtX gene), respectively. The

resulting plasmid, pAC-BETA, retains functional genes for geranylgeranyl pyrophosphate synthase (crtE), phytoene synthase (crtB), phytoene desaturase (crtI), and lycopene cyclase (crtY). Cells of *E. coli* containing this plasmid form yellow colonies and accumulate β-carotene. A plasmid containing both the lycopene ϵ- and β-cyclase cDNAs of *A. thaliana* was constructed by excising the ϵ-cyclase in clone y2 as a PvuI-PvuII fragment and ligating this piece in the SnaBI site of a plasmid (pSPORT 1 from GIBCO-BRL) that already contained the β-cyclase (Cunningham et al., 1996).

Organisms and Growth Conditions

E. coli strains TOP10 and TOP10 F' (obtained from Invitrogen Corporation, San Diego, CA) and XL1-Blue (Stratagene) were grown in Luria-Bertani (LB) medium (Sambrook et al., 1989) at 37°C in darkness on a platform shaker at 225 cycles per min. Media components were from Difco (yeast extract and tryptone) or Sigma (NaCl). Ampicillin at 150 μg/mL and/or chloramphenicol at 50 μg/mL (both from United States Biochemical Corporation) were used, as appropriate, for selection and maintenance of plasmids.

Mass Excision and Color Complementation Screening of an *A. thaliana* cDNA Library

A size-fractionated 1-2 kB cDNA library of *A. thaliana* in lambda ZAPII (Kieber et al., 1993) was obtained from the Arabidopsis Biological Resource Center at The Ohio State University (stock number CD4-14). Other size fractionated libraries were also obtained (stock numbers CD4-13, CD4-15, and CD4-16). An aliquot of each library was treated to cause a mass excision of the cDNAs and thereby produce a phagemid library according to the instructions provided by the supplier of the cloning vector (Stratagene; *E. coli* strain XL1-Blue and the helper phage R408 were used). The titre of the excised phagemid was determined and the library was introduced into a lycopene-accumulating strain of *E. coli* TOP10 F' (this strain contained the plasmid pAC-LYC) by incubation of the phagemid with the *E. coli* cells for 15 min at 37°C. Cells had been grown overnight at 30°C in LB medium supplemented with 2% (w/v) maltose and 10 mM $MgSO_4$ (final concentration), and harvested in 1.5 ml microfuge tubes at a setting of 3 on an Eppendorf microfuge (5415C) for 10 min. The pellets were resuspended in 10 mM $MgSO_4$ to a volume equal to one-half that of the

initial culture volume. Transformants were spread on large (150 mm diameter) LB agar petri plates containing antibiotics to provide for selection of cDNA clones (ampicillin) and maintenance of pAC-LYC (chloramphenicol). Approximately 10,000 colony forming units were spread on each plate. Petri plates were incubated at 37·C for 16 hr and then at room temperature for 2 to 7 days to allow maximum color development. Plates were screened visually with the aid of an illuminated 3x magnifier and a low power stage-dissecting microscope for the rare, pale pinkish-yellow to deep-yellow colonies that could be observed in the background of pink colonies. A colony color of yellow or pinkish-yellow was taken as presumptive evidence of a cyclization activity. These yellow colonies were collected with sterile toothpicks and used to inoculate 3ml of LB medium in culture tubes with overnight growth at 37°C and shaking at 225 cycles/min. Cultures were split into two aliquots in microfuge tubes and harvested by centrifugation at a setting of 5 in an Eppendorf 5415C microfuge. After discarding the liquid, one pellet was frozen for later purification of plasmid DNA. To the second pellet was added 1.5 ml EtOH, and the pellet was resuspended by vortex mixing, and extraction was allowed to proceed in the dark for 15-30 min with occasional remixing. Insoluble materials were pelleted by centrifugation at maximum speed for 10 min in a microfuge. Absorption spectra of the supernatant fluids were recorded from 350-550 nm with a Perkin Elmer lambda six spectrophotometer.

Analysis of isolated clones

Eight of the yellow colonies contained β-carotene indicating that a single gene product catalyzes both cyclizations required to form the two β endgroups of the symmetrical β-carotene from the symmetrical precursor lycopene. One of the yellow colonies contained a pigment with the spectrum characteristic of δ-carotene, a monocyclic carotenoid with a single ε endgroup. Unlike the β cyclase, this ε-cyclase appears unable to carry out a second cyclization at the other end of the molecule.

The observation that ε-cyclase is unable to form two cyclic ε-endgroups (*e.g.* the bicyclic ε-carotene) illuminates the mechanism by which plants can coordinate and control the flow of substrate into carotenoids derived from β-carotene versus those derived from α-carotene and also can prevent the formation of carotenoids with two ε endgroups.

The availability of the *A. thaliana* gene encoding the ε-cyclase enables the directed manipulation of plant and algal species for modification of carotenoid content and

composition. Through inactivation of the ϵ-cyclase, whether at the gene level by deletion of the gene or by insertional inactivation or by reduction of the amount of enzyme formed (by such as antisense technology), one may increase the formation of β-carotene and other pigments derived from it. Since vitamin A is derived only from carotenoids with β endgroups, an enhancement of the production of β-carotene versus α-carotene may enhance nutritional value of crop plants. Reduction of carotenoids with ϵ-endgroups may also be of value in modifying the color properties of crop plants and specific tissues of these plants. Alternatively, where production of α-carotene, or pigments such as lutein that are derived from α-carotene, is desirable, whether for the color properties, nutritional value or other reason, one may overexpress the ϵ-cyclase or express it in specific tissues. Wherever agronomic value of a crop is related to pigmentation provided by carotenoid pigments the directed manipulation of expression of the ϵ-cyclase gene and/or production of the enzyme may be of commercial value.

The predicted amino acid sequence of the *A. thaliana* ϵ-cyclase enzyme was determined. A comparison of the amino acid sequences of the β- and ϵ-cyclase enzymes of *Arabidopsis thaliana* (Fig. 13) as predicted by the DNA sequence of the respective cDNAs (Fig. 4 for the ϵ-cyclase cDNA sequence), indicates that these two enzymes have many regions of sequence similarity, but they are only about 37% identical overall at the amino acid level. The degree of sequence identity at the DNA base level, only about 50%, is sufficiently low such that we and others have been unable to detect this gene by hybridization using the β cyclase as a probe in DNA gel blot experiments.

REFERENCES

Each reference cited in this application and/or listed below is hereby incorporated by reference.

Bird et al, 1991 Biotechnology 9, 635-639.

Bishop et al., (1995) FEBS Lett. 367, 158-162.

Bramley, P.M. (1985) Adv. Lipid Res. 21, 243-279.

Bramley, P.M. (1992) Plant J. 2, 343-349.

Britton, G. (1988). Biosynthesis of carotenoids. In Plant Pigments, T.W. Goodwin, ed. (London: Academic Press), pp. 133-182.

Britton, G. (1979) Z. Naturforsch. Section C Biosci. 34, 979-985.

Britton, G. (1995) UV/Visible spectroscopy. In Carotenoids, Vol. IB: Spectroscopy, G. Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 13-62.

Bouvier et al., (1994) Plant J. 6, 45-54.

Cunningham et al., (1985) Photochem. Photobiol. 42: 295-307.

Cunningham et al., (1993) FEBS Lett. 328, 130-138.

Cunningham et al., (1994) Plant Cell 6, 1107-1121.

Cunningham et al., (1996) Plant Cell 8, 1613-1626.

Davies, B.H. (1976). Carotenoids. In Chemistry and Biochemistry of Plant Pigments, Vol. 2, T.W. Goodwin, ed (New York: Academic Press), pp. 38-165.

Del Sal et al., (1988). Nucl. Acids Res. 16, 9878.

Demmig-Adams & Adams, (1992) Ann. Rev. Plant Physiol. Mol. Biol. 43, 599-626.

Enzell & Back, (1995) Mass spectrometry. In Carotenoids, Vol. IB: Spectroscopy, G. Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 261-320.

Frank & Cogdell (1993) Photochemistry and function of carotenoids in photosynthesis. In Carotenoids in Photosynthesis. A. Young and G. Britton, eds. (London: Chapman and Hall). pp. 253-326.

Goodwin, T.W. (1980). The Biochemistry of the Carotenoids. 2nd ed, Vol. 1 (London: Chapman and Hall.

Horvath et al., (1972) Phytochem. 11, 183-187.

Hugueney et al., (1995) Plant J. 8, 417-424.

Hundle et al., (1991) Photochem. Photobiol. 54, 89-93.

Jensen & Jensen, (1971) Methods Enzymol. 23, 586-602.

Kargl & Quackenbush, (1960) Archives Biochem. Biophys. 88, 59-63.

Kargl et al., (1960) Proc. Am. Hort. Soc. 75, 574-578.

Kieber et al., (1993) Cell 72, 427-441.

Koyama, Y. (1991) J. Photochem. Photobiol., B, 9, 265-80.

Krinsky, N.I. (1987) Medical uses of carotenoids. In Carotenoids, N.I. Krinsky, M.M. Mathews-Roth, and R.F. Taylor, eds. (New York: Plenum), pp. 195-206.

Kyte & Doolittle, (1982) J. Mol. Biol. 157, 105-132.

LaRossa & Schloss, (1984) J. Biol. Chem. 259, 8753-8757.

Misawa et al., (1994a) Plant J. 6, 481-489.

Misawa et al., (1994b) J. Biochem, Tokyo, 116, 980-985.

Norris et al., (1995) Plant Cell 7, 2139-2149.

Pecker et al., (1996) Submitted to Plant Mol. Biol.

Perry et al., (1986) J. Bacteriol. 168, 607-612.

Persson & Argos, (1994) J. Mol. Biol. 237, 182-192.

Plumley & Schmidt, (1987) Proc. Nat. Acad. Sci. USA 83, 146-150.

Plumley & Schmidt, (1995) Plant Cell 7, 689-704.

Rossmann et al., (1974) Nature 250, 194-199.

Rock & Zeevaart (1991) Proc. Nat. Acad. Sci. USA 88, 7496-7499.

Rost et al., (1995) Protein Science 4, 521-533.

Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, 2nd edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Sancar, A. (1994) Biochemistry 33, 2-9.

Sander & Schneider, (1991) Proteins 9, 56-68.

Sandmann, G. (1994) Eur. J. Biochem. 223, 7-24.

Scolnik & Bartley, (1995) Plant Physiol. 108, 1342.

Siefermann-Harms, D. (1987) Physiol. Plant. 69, 561-568.

Spurgeon & Porter, (1980). Biosynthesis of carotenoids. In Biochemistry of Isoprenoid Compounds, J.W. Porter, and S.L. Spurgeon, eds. (New York: Wiley), pp. 1-122.

Tomes, M.L. (1963) Bot. Gaz. 124, 180-185.

Tomes, M.L. (1967) Genetics 56, 227-232.

Tuveson et al., (1986) J. Bacteriol. 170, 4675-4680.

Van Beeumen et al., (1991) J. Biol. Chem. 266, 12921-12931.

Weedon & Moss, (1995) Structure and Nomenclature. In Carotenoids, Vol. IB: Spectroscopy, G. Britton, S. Liaaen-Jensen, H.P. Pfander, eds. (Basel: Birkhauser Verlag), pp. 27-70.

Wierenga et al., (1986) J. Mol. Biol. 187, 101-107.

Zechmeister, L. (1962) Cis-Trans Isomeric Carotenoids, Vitamins A and Arylpolyenes. Springer-Verlag, Vienna.

Having now fully described the invention, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the invention as set forth herein.

We claim:

1. An isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ-cyclase enzyme activity and has an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23 or 25-27.

2. The nucleic acid sequence of claim 1, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23 or 25-27.

3. A vector comprising the nucleic acid sequence of claim 1, wherein the nucleic acid sequence is operably linked to a promoter.

4. A host cell which contains the vector of claim 3.

5. The host cell of claim 4, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell, a yeast cell and a plant cell.

6. The host cell of claim 4, wherein the host cell is a photosynthetic cell.

7. An isolated and/or purified protein having lycopene ϵ-cyclase enzyme activity and having an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23 or 25-27.

8. The protein of claim 7, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23 or 25-27.

AMENDED CLAIMS

[received by the International Bureau on 15 November 1999 (15.11.99);
original claims 1,2,7 and 8 amended; remaining claims unchanged (1 page)]

1. An isolated and/or purified nucleic acid sequence which encodes for a protein having lycopene ϵ-cyclase enzyme activity and has an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23, 25 or 26.

2. The nucleic acid sequence of claim 1, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23, 25 or 26.

3. A vector comprising the nucleic acid sequence of claim 1, wherein the nucleic acid sequence is operably linked to a promoter.

4. A host cell which contains the vector of claim 3.

5. The host cell of claim 4, wherein the host cell is selected from the group consisting of a bacterial cell, an algal cell, a yeast cell and a plant cell.

6. The host cell of claim 4, wherein the host cell is a photosynthetic cell.

7. An isolated and/or purified protein having lycopene ϵ-cyclase enzyme activity and having an amino acid sequence which is at least 85% identical to one of SEQ ID NOS: 23, 25 or 26.

8. The protein of claim 7, wherein the protein has the amino acid sequence of one of SEQ ID NOS: 23, 25 or 26.

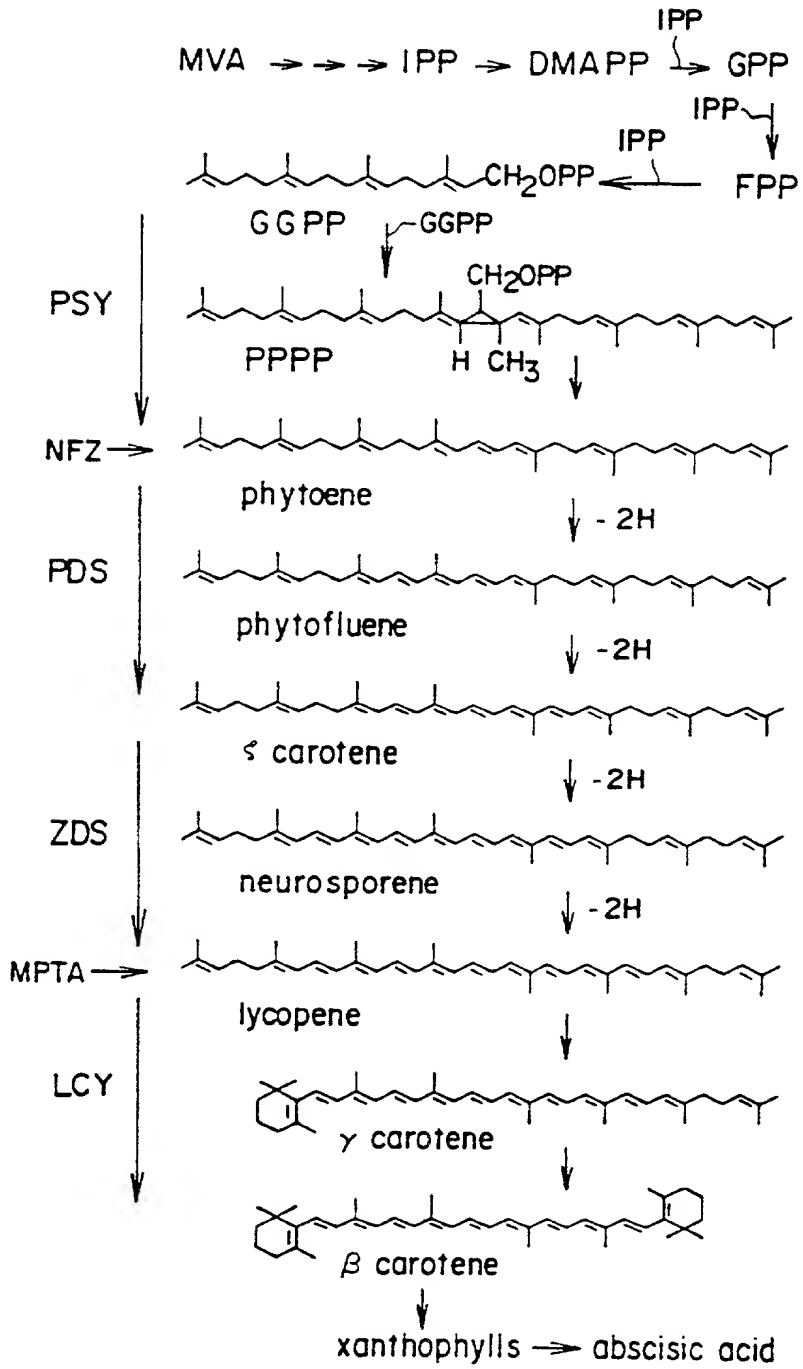


FIG. 1

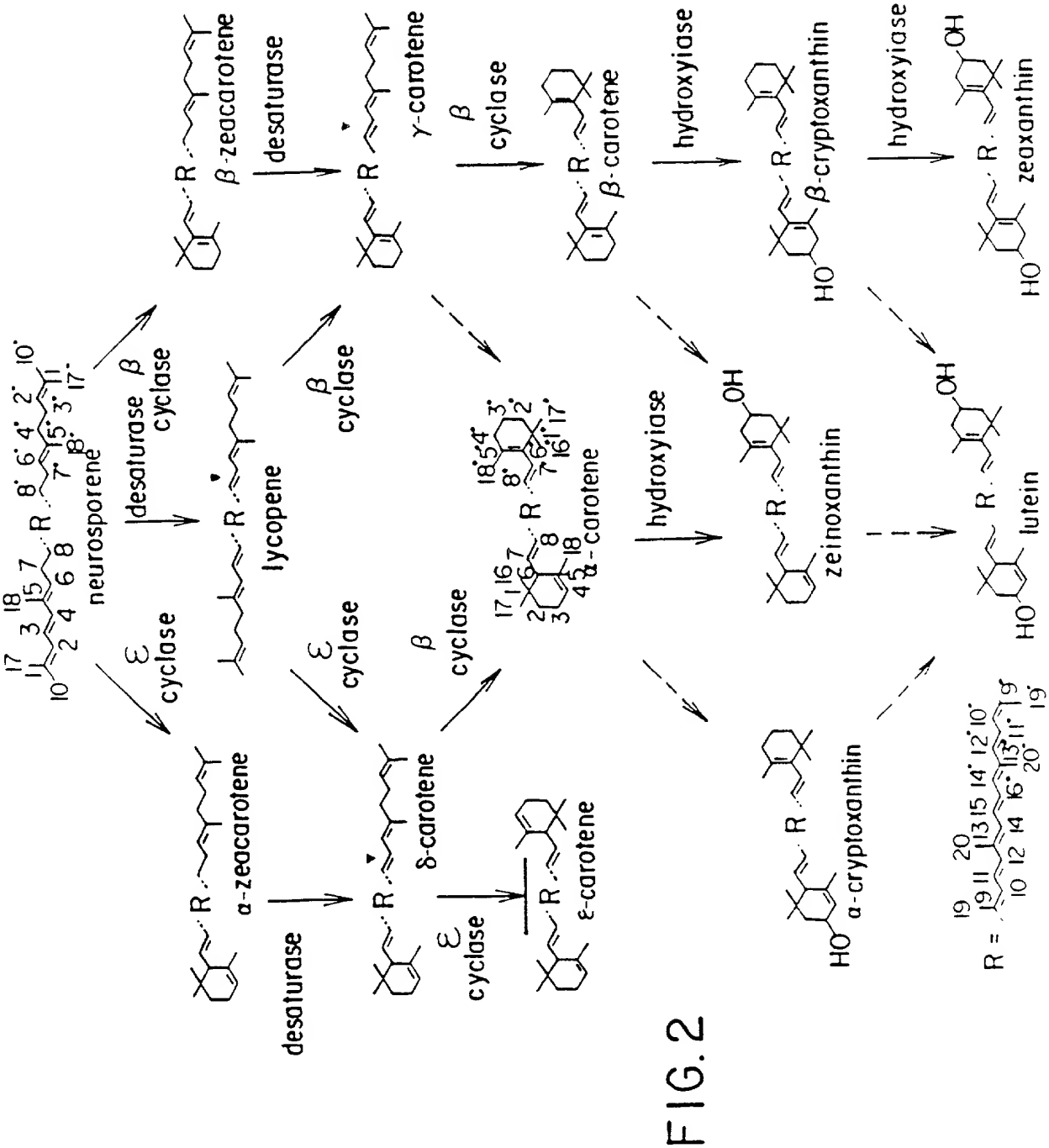


FIG. 2

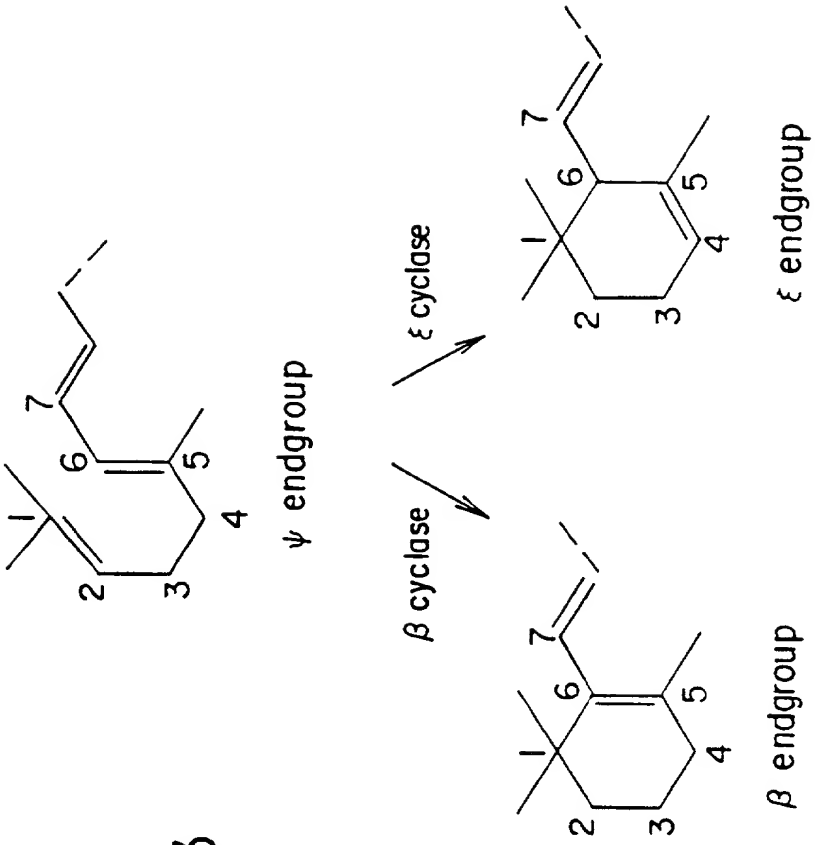


FIG. 3

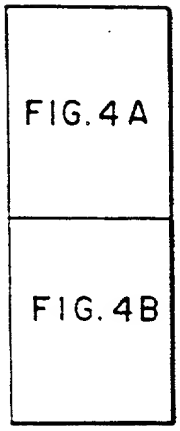


FIG. 4

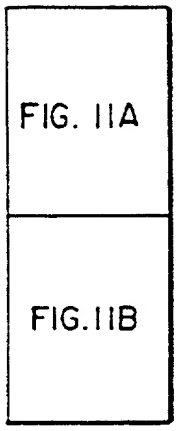


FIG. 11

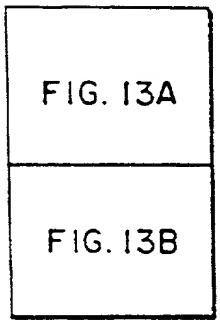


FIG. 13

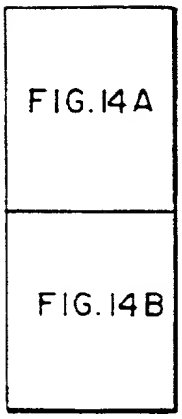


FIG. 14

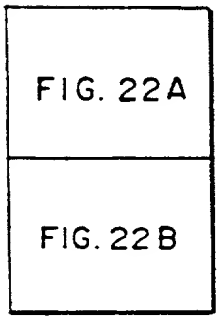


FIG. 22

FIG. 4A

Arabidopsis thaliana epsilon cyclase:

```
                                                                    acaaaaggaaataattag attcctctttctgcttgctataccttgaca   48

                                                gaacaacataacaatggtgtaagtcttctc gctgtattcgaaattatttggaggaggaac  108

                                                atggagtgtgttggggctaggaatttcgca gcaatggcggtttcaacatttccgtcatgg  168
                                             1  M  E  C  V  G  A  R  N  F  A  A  M  A  V  S  T  F  P  S  W

                                                agttgtcgaaggaaatttccagtggctaag agatacagctataggaatattcgcttcggt  228
                                            21  S  C  R  R  K  F  P  V  V  K  R  Y  S  Y  R  N  I  R  F  G

                                                ttgtgtagtgtcagagctagcggcggcgga agttccggtagtgagagttgtgtagcggtg  288
                                            41  L  C  S  V  R  A  S  G  G  G  S  S  G  S  K  S  C  V  A  V

                                                agagaagatttcgctgacgaagaagatttt gcgaaagctggcggttctgagattctattt  348
                                            61  R  S  D  F  A  D  E  E  D  F  V  E  A  G  G  S  R  I  L  F

                                                gttcaaatgcagcagaacaaagatatggat gaacagtctaagcttgttgataagttgcct  408
                                            81  V  Q  M  Q  Q  M  K  D  M  D  S  Q  S  K  L  V  D  K  L  P

                                                cctatatcaactggtgatggtgctttggat catgtggttactggctgtggtcctgctggt  468
                                           101  P  I  S  I  G  D  G  A  L  D  K  V  V  I  G  C  G  P  A  G

                                                ttagccttggctgcagaatcagctaagctt ggattaaaagttggactcattggtccagat  528
                                           121  L  A  L  A  A  K  S  A  K  L  G  L  K  V  G  L  I  G  P  D

                                                cttccttttactaacaattacggtgtttgg gaagatgaattcaatgatcttgggctgcaa  588
                                           141  L  P  F  T  M  M  Y  G  V  M  K  D  K  F  N  D  L  G  L  G

                                                aaatgtattgagcatgtttggagagagact attgcgcacctggatgatgacaagcctatt  648
                                           161  K  C  I  K  K  V  W  R  S  T  I  V  Y  L  D  D  D  K  P  I

                                                accattggccgtgcttatggaagagttagt cgacgtttgctccatgaggagcttttgagg  708
                                           181  T  I  G  R  A  Y  G  R  V  S  R  R  L  L  X  E  E  L  L  R

                                                aggtgtgtcgagtcaggtgtctcgtaccttt agctcgaaagttgacagcataacagaagct  768
                                           201  R  C  V  K  S  G  V  S  Y  L  S  S  K  V  D  S  I  T  E  A

                                                tgtgatggccttagacttgttgcttgtgac gacaataacgtcattccctgcaggcttgcc  828
                                           221  S  D  G  L  X  L  V  A  C  D  D  M  M  V  I  P  C  X  L  A

                                                actgttgcttctggagcagcttcgggaaag ctcttgcaatacgaagttggtggacctaga  888
                                           241  T  V  A  S  G  A  A  S  G  K  L  L  Q  Y  X  V  G  G  P  R

                                                gtctgtgcgcaaactgcatacggcgtggag gttgaggcggaaaatagtccatatgatcca  948
```

FIG. 4B

```
261  V C V Q T A Y G V X  V X V X N S P Y D P

     gatcaaatggttttcatggattacagagat tatactaacgagaaagttcggagcttagaa 1008
281  D Q M V P M D Y R D  Y T M X X V R S L X

     gctgagtatccaacgtttctgtacgccatg cctatgacaaagtcaagactcttcttcgag 1068
301  A K Y P T F L Y A M  P M T K S R L F F K

     gagacatgtttggcctcaaaagatgtcatg ccctttgatttgctaaaaacgaagctcatg 1128
321  K T C L A S K D V M  P F D L L K T K L M

     ttaagattagacacactcggaattcgaatt ctaaagacttacgaagaggagtggtcctat 1188
341  I P V G G S L P N T  X Q K N L A F G A A

     atcccagttggtggttccttgccaaacacc gaacaaaagaatctcgcctttggtgctgcc 1248
361  I P V G G S L P M T  X Q K N L A F G A A

     gctagcatggtacatcccgcaacaggctat tcagttgtgagatctttgtctgaagctcca 1308
381  A S M V M P A T G Y  S V V R S L S X A P

     aaacatgcatcagtcatcgcagagatacta agagaagagactaccaaacagattaacagt 1368
401  K Y A S V I A K I L  R E E T T K Q I N S

     aatatttcaagacaagcttaggatacttta tggccaccagaaaggaaaagacagagagca 1428
421  M I S R Q A W D T L  W P P E R X R Q R A

     ttctttctctttggtcttgcactcagagtt caattcgataccgaaggcattagaagcttc 1488
441  F F L F G L A L I V  Q F D T X G I R S F

     ttccgtactttcttccgccttccaaaatgg atgtggcaagggtttctaggatcaacatta 1548
461  F R T P F R L P K W  M W Q G F L G S T L

     acatcaggagatctcgttctctttgcttta tacatgttcgtcatttcaccaaacaatttg 1608
481  T S G D L V L F A L  Y M P V I S P M M L

     agaaaaggtctcattaatcatctcatctct gatccaaccggagcaaccatgataaaaacc 1668
501  R K G L I N W L I S  D P T G A T M I K T

     tatctcaaagtatgatttacttaccaactc ttaggtttgtgtatatatatgccgatttat 1728

521  Y L K V

     ctgaataatcgatcaaagaatggtatgtgg gttactaggaagttggaaacaaacacgtat 1788

     agaatctaaggagtgatcgaaatggagacg gaaacgaaaagaaaaaaatcagtctttgtt 1848

     ccgtggctagtg                                                  1868
```

FIG. 5

```
   1  gctctttctc ctcctcctct accgatttcc gactccgcct cccgaaatcc
  51  ttatccggat tctctccgtc tcttcgattt aaacgctttt ctgtctgtta
 101  cgtcgtcgaa gaacggagac agaattctcc gattgagaac gatgagagac
 151  cggagagcac gagctccaca aacgctatag acgctgagta tctggcgttg
 201  cgtttggcgg agaaattgga gaggaagaaa tcggagaggt ccacttatct
 251  aatcgctgct atgttgtcga gctttggtat cacttctatg gctgttatgg
 301  ctgtttacta cagattctct tggcaaatgg agggaggtga gatctcaatg
 351  ttggaaatgt ttggtacatt tgctctctct gttggtgctg ctgttggtat
 401  ggaattctgg gcaagatggg ctcatagagc tctgtggcac gcttctctat
 451  ggaatatgca tgagtcacat cacaaaccaa gagaaggacc gtttgagcta
 501  aacgatgttt ttgctatagt gaacgctggt ccagcgattg gtctcctctc
 551  ttatggattc ttcaataaag gactcgttcc tggtctctgc tttggcgccg
 601  ggttaggcat aacggtgttt ggaatcgcct acatgtttgt ccacgatggt
 651  ctcgtgcaca agcgtttccc tgtaggtccc atcgccgacg tcccttacct
 701  ccgaaaggtc gccgccgctc accagctaca tcacacagac aagttcaatg
 751  gtgtaccata tggactgttt cttggaccca aggaattgga agaagttgga
 801  ggaaatgaag agttagataa ggagattagt cggagaatca aatcatacaa
 851  aaaggcctcg ggctccgggt cgagttcgag ttcttgactt taaacaagtt
 901  ttaaatccca aattcttttt ttgtcttctg tcattatgat catcttaaga
 951  cggtct
```

FIG. 6

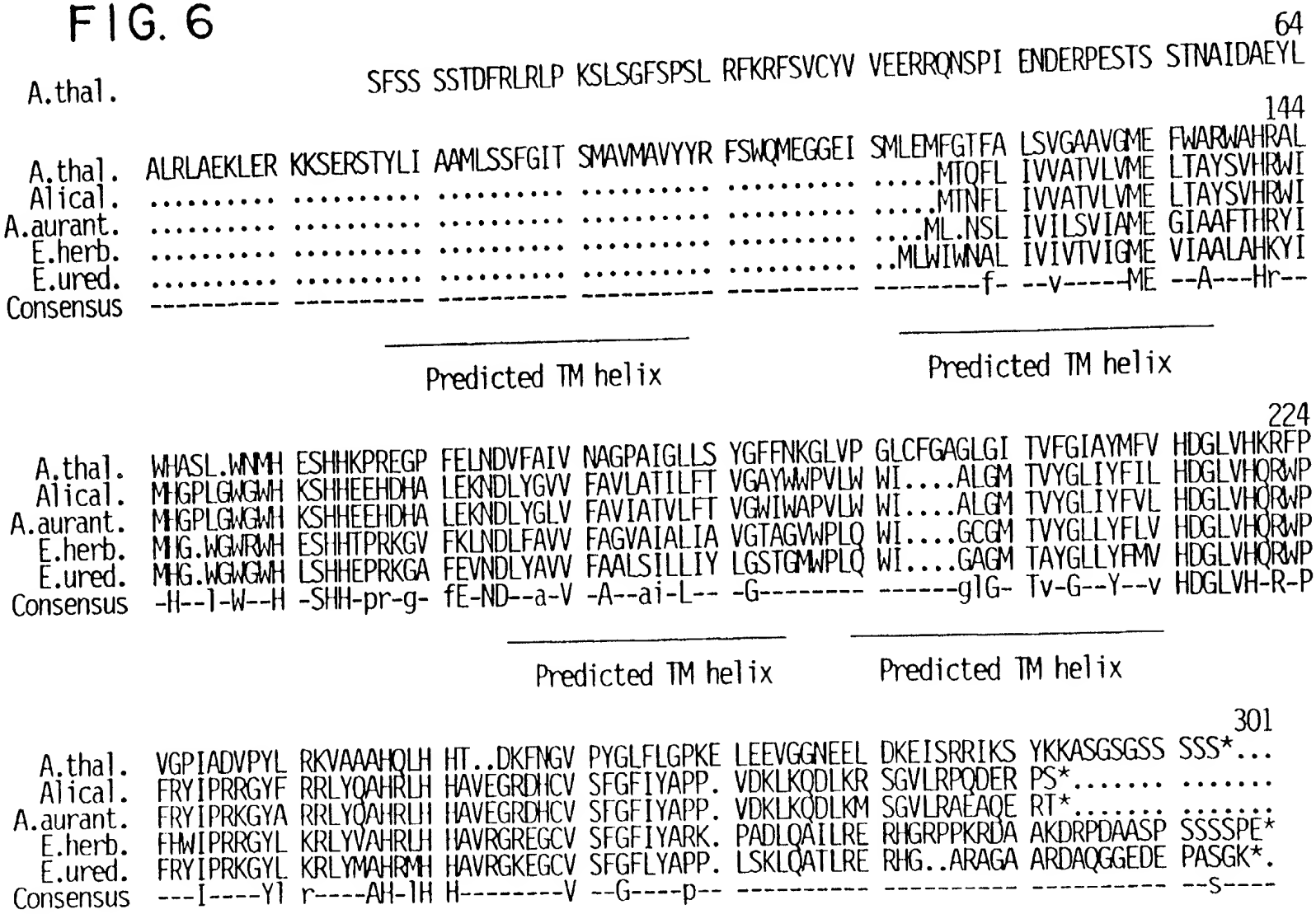


FIG. 7

```
   1  ccacgggtcc gcctcccccgt ttttttccga tccgatctcc ggtgccgagg
  51  actcagctgt ttgttcgcgc tttctcagcc gtcaccatga ccgattctaa
 101  cgatgctgga atggatgctg ttcagagacg actcatgttt gaagacgaat
 151  gcattctcgt tgatgaaaat aatcgtgtgg tgggacatga cactaagtat
 201  aactgtcatc tgatggaaaa gattgaagct gagaatttac ttcacagagc
 251  tttcagtgtg tttttattca actccaagta tgagttgctt ctccagcaac
 301  ggtcaaaaac aaaggttact ttcccacttg tgtggacaaa cacttgttgc
 351  agccatcctc tttaccgtga atccgagctt attgaagaga atgtgcttgg
 401  tgtaagaaat gccgcacaaa ggaagctttt cgatgagctc ggtattgtag
 451  cagaagatgt accagtcgat gagttcactc ccttgggacg catgctttac
 501  aaggcacctt ctgatgggaa atggggagag cacgaagttg actatctact
 551  cttcatcgtg cgggatgtga agcttcaacc aaacccagat gaagtggctg
 601  agatcaagta cgtgagcagg gaagagctta aggagctggt gaagaaagca
 651  gatgctggcg atgaagctgt gaaactatct ccatggttca gattggtggt
 701  ggataatttc ttgatgaagt ggtgggatca tgttgagaaa ggaactatca
 751  ctgaagctgc agacatgaaa accattcaca agctctgaac tttccataag
 801  ttttggatct tccccttccc ataataaaat taagagatga gacttttatt
 851  gattacagac aaaactggca acaaaatcta ttcctaggat ttttttttgc
 901  tttttattta cttttgattc atctctagtt tagttttcat cttaaaaaaa
 951  aaaa
```

FIG. 8

```
   1  caccaatgtc tgtttcttct ttatttaatc tcccattgat tcgcctcaga
  51  tctctcgctc tttcgtcttc tttttcttct ttccGATTTG CCCATCGTCC
 101  TCTGTCATCG ATTTCACCGA GAAAGTTACC GAATTTTCGT GCTTTCTCTG
 151  GTACCGCTAT GACAGATACT AAAGATGCTG GTATGGATGC TGTTCAGAGA
 201  CGTCTCATGT TTGAGGATGA ATGCATTCTT GTTGATGAAA CTGATCGTGT
 251  TGTGGGGCAT GTCAGCAAGT ATAATTGTCA TCTGATGGAA AATATTGAAG
 301  CCAAGAATTT GCTGCACAGG GCTTTTAGTG TATTTTTATT CAACTCGAAG
 351  TATGAGTTGC TTCTCCAGCA AAGGTCAAAC ACAAAGGTTA CGTTCCCTCT
 401  AGTGTGGACT AACACTTGTT GCAGCCATCC TCTTTACCGT GAATCAGAGC
 451  TTATCCAGGA CAATGCACTA GGTGTGAGGA ATGCTGCACA AAGAAAGCTT
 501  CTCGATGAGC TTGGTATTGT AGCTGAAGAT GTACCAGTCG ATGAGTTCAC
 551  TCCCTTGGGA CGTATGCTGT ACAAGGCTCC TTCTGATGGC AAATGGGGAG
 601  AGCATGAACT TGATTACTTG CTCTTCATCG TGCGAGACGT GAAGGTTCAA
 651  CCAAACCCAG ATGAAGTAGC TGAGATCAAG TATGTGAGCC GGGAAGAGCT
 701  GAAGGAGCTG GTGAAGAAAG CAGATGCAGG TGAGGAAGGT TTGAAACTGT
 751  CACCATGGTT CAGATTGGTG GTGGACAATT TCTTGATGAA GTGGTGGGAT
 801  CATGTTGAGA AAGGAACTTT GGTTGAAGCT ATAGACATGA AAACCATCCA
 851  CAAACTCTGA ACATCTTTTT TTAAAGTTTT TAAATCAATC AACTTTCTCT
 901  TCATCATTTT TATCTTTTCG ATGATAATAA TTTGGGATAT GTGAGACACT
 951  TACAAAACTT CCAAGCACCT CAGGCAATAA TAAAGTTTGC GGCCGC
```

FIG. 9

```
   1  CTCGGTAGCT GGCCACAATC GCTATTTGGA ACCTGGCCCG GCGGCAGTCC
  51  GATGCCGCGA TGCTTCGTTC GTTGCTCAGA GGCCTCACGC ATATCCCCCG
 101  CGTGAACTCC GCCCAGCAGC CCAGCTGTGC ACACGCGCGA CTCCAGTTTA
 151  AGCTCAGGAG CATGCAGATG ACGCTCATGC AGCCCAGCAT CTCAGCCAAT
 201  CTGTCGCGCG CCGAGGACCG CACAGACCAC ATGAGGGGTG CAAGCACCTG
 251  GGCAGGCGGG CAGTCGCAGG ATGAGCTGAT GCTGAAGGAC GAGTGCATCT
 301  TGGTGGATGT TGAGGACAAC ATCACAGGCC ATGCCAGCAA GCTGGAGTGT
 351  CACAAGTTCC TACCACATCA GCCTGCAGGC CTGCTGCACC GGGCCTTCTC
 401  TGTGTTCCTG TTTGACGATC AGGGGCGACT GCTGCTGCAA CAGCGTGCAC
 451  GCTCAAAAAT CACCTTCCCA AGTGTGTGGA CGAACACCTG CTGCAGCCAC
 501  CCTTTACATG GGCAGACCCC AGATGAGGTG GACCAACTAA GCCAGGTGGC
 551  CGACGGAACA GTACCTGGCG CAAAGGCTGC TGCCATCCGC AAGTTGGAGC
 601  ACGAGCTGGG GATACCAGCG CACCAGCTGC CGGCAAGCGC GTTTCGCTTC
 651  CTCACGCGTT TGCACTACTG TGCCGCGGAC GTGCAGCCAG CTGCGACACA
 701  ATCAGCGCTC TGGGGCGAGC ACGAAATGGA CTACATCTTG TTCATCCGGG
 751  CCAACGTCAC CTTGGCGCCC AACCCTGACG AGGTGGACGA AGTCAGGTAC
 801  GTGACGCAAG AGGAGCTGCG GCAGATGATG CAGCCGGACA ACGGGCTGCA
 851  ATGGTCGCCG TGGTTTCGCA TCATCGCCGC GCGCTTCCTT GAGCGTTGGT
 901  GGGCTGACCT GGACGCGGCC CTAAACACTG ACAAACACGA GGATTGGGGA
 951  ACGGTGCATC ACATCAACGA AGCGTGAAAG CAGAAGCTGC AGGATGTGAA
1001  GACACGTCAT GGGGTGGAAT TGCGTACTTG GCAGCTTCGT ATCTCCTTTT
1051  TCTGAGACTG AACCTGCAGT CAGGTCCCAC AAGGTCAGGT AAAATGGCTC
1101  GATAAAATGT ACCGTCACTT TTTGTCGCGT ATACTGAACT CCAAGAGGTC
1151  AAAAAAAAAA AAAAA
```

09/701395

FIG. 10

```
1     CTCGGTAGCT GGCCACAATC GCTATTTGGA ACCTGGCCCG GCGGCAGTCC
51    GATGCCGCGA TGCTTCGTTC GTTGCTCAGA GGCCTCACGC ATATCCCGCG
101   CGTGAACTCC GCCCAGCAGC CCAGCTGTGC ACACGCGCGA CTCCAGTTTA
151   AGCTCAGGAG CATGCAGCTG CTTTCCGAGG ACCGCACAGA CCACATGAGG
201   GGTGCAAGCA CCTGGGCAGG CGGGCAGTCG CAGGATGAGC TGATGCTGAA
251   GGACGAGTGC ATCTTGGTAG ATGTTGAGGA CAACATCACA GGCCATGCCA
301   GCAAGCTGGA GTGTCACAAG TTCCTACCAC ATCAGCCTGC AGGCCTGCTG
351   CACCGGGCCT TCTCTGTGTT CCTGTTTGAC GATCAGGGGC GACTGCTGCT
401   GCAACAGCGT GCACGCTCAA AAATCACCTT CCCAAGTGTG TGGACGAACA
451   CCTGCTGCAG CCACCCTTTA CATGGGCAGA CCCCAGATGA GGTGGACCAA
501   CTAAGCCAGG TGGCCGACGG AACAGTACCT GGCGCAAAGG CTGCTGCCAT
551   CCGCAAGTTG GAGCACGAGC TGGGGATACC AGCGCACCAG CTGCCGGCAA
601   GCGCGTTTCG CTTCCTCACG CGTTTGCACT ACTGTGCCGC GGACGTGCAG
651   CCAGCTGCGA CACAATCAGC GCTCTGGGGC GAGCACGAAA TGGACTACAT
701   CTTGTTCATC CGGGCCAACG TCACCTTGGC GCCCAACCCT GACGAGGTGG
751   ACGAAGTCAG GTACGTGACG CAAGAGGAGC TGCGGCAGAT GATGCAGCCG
801   GACAACGGGC TTCAATGGTC GCCGTGGTTT CGCATCATCG CCGCGCGCTT
851   CCTTGAGCGT TGGTGGGCTG ACCTGGACGC GGCCCTAAAC ACTGACAAAC
901   ACGAGGATTG GGGAACGGTG CATCACATCA ACGAAGCGTG AAGGCAGAAG
951   CTGCAGGATG TGAAGACACG TCATGGGGTG GAATTGCGTA CTTGGCAGCT
1001  TCGTATCTCC TTTTTCTGAG ACTGAACCTG CAGAGCTAGA GTCAATGGTG
1051  CATCATATTC ATCGTCTCTC TTTTGTTTTA GACTAATCTG TAGCTAGAGT
1101  CACTGATGAA TCCTTTACAA CTTTCAAAAA AAAAA
```

09/701595

FIG. 11A

```
          1                                                  50
   HPO4 MLRSLLRGLT HIPRVNSAQQ PSCAHARLQF KLRSMQMTLM QPSISANLSR
   HPO5 MLRSLLRGLT HIPRVNSAQQ PSCAHARLQF KLRSMQLL.. ..........
   ATDP7 MSVSSLFNLP .LIRLRSLA. LSSSFSSFRF AHRPLSSIS. PRKLPNFRAF
 C.brew. MS.SSMLNFT .ASRIVSLPL LSSPPSRVHL PLCFFSPISL TQRFSAKLTF
   ATOP5 .......... .TGPPPRFFP IRSPVPRTQL FVRAFSAV.. ..........
 S.cerev. ..MTADNNSM PHGAVSSYAK LVQNQTPEDI LEEFPEIIPL QQRPN...TR

          51                                                 100
   AEDRTDHMRG ASTWAGGQSQ DELMLKDECI LVDVEDNITG HASKLECHKF
   SEDRTDHMRG ASTWAGGQSQ DELMLKDECI LVDVEDNITG HASKLECHKF
   S..GTA.MTD TKDAGMDAVQ RRLMFEDECI LVDETDRVVG HVSKYNCHLM
   SSQATT.MGE VVDAGMDAVQ RRLMFEDECI LVDENDKVVG HESKYNCHLM
   .....T.MTD SNDAGMDAVQ RRLMFEDECI LVDENNRVVG HDTKYNCHLM
   SSETSNDESG ETCFSGHDEE QIKLMNENCI VLDWDDNAIG AGTKKVCHLM

          101                                                150
   LPHQPAGLLH RAFSVFLFDD QGRLLLQQRA RSKITFPSVW TNTCCSHPLH
   LPHQPAGLLH RAFSVFLFDD QGRLLLQQRA RSKITFPSVW TNTCCSHPLH
   ENIEAKNLLH RAFSVFLFNS KYELLLQQRS NTKVTFPLVW TNTCCSHPLY
   ENIESENLLH RAFSVFLFNS KYELLLQQRS ATKVTFPLVW TNTCCSHPLY
   EKIEAENLLH RAFSVFLFNS KYELLLQQRS KTKVTFPLVW TNTCCSHPLY
   ENIE.KGLLH RAFSVFIFNE QGELLLQQRA TEKITFPDLW TNTCCSHPLC

          151                                                200
   GQTPDEVDQL SQVADGTVPG AKAAAIRKLE HELGIPAHQL PA.SAFRFLT
   GQTPDEVDQL SQVADGTVPG AKAAAIRKLE HELGIPAHQL PA.SAFRFLT
   RE..... SELIQDNALG VRNAAQRKLL DELGIVAEDV PV.DEFTPLG
   RE..... SELIDENCLG VRNAAQRKLL DELGIPAEDL PV.DQFIPLS
   RE..... SELIEENVLG VRNAAQRKLF DELGIVAEDV PV.DEFTPLG
   ID...DELGL KGKLDDKIKG AITAAVRKLD HELGIPEDET KTRGKFHFLN

          201                                                250
   RLHYCAADVQ PAATQSALWG EHEMDYILFI ....RANVTL APNPDEVDEV
   RLHYCAADVQ PAATQSALWG EHEMDYILFI ....RANVTL APNPDEVDEV
   RMLY..... .KAPSDGKWG EHELDYLLFI ....VRDVKV QPNPDEVAEI
   RILY..... .KAPSDGKWG EHELDYLLFI ....IRDVNL DPNPDEVAEV
   RMLY..... .KAPSDGKWG EHEVDYLLFI ....VRDVKL QPNPDEVAEI
   RIHY..... .MAPSNEPWG EHEIDYILFY KINAKENLTV NPNVNEVRDF
```

FIG. 11B

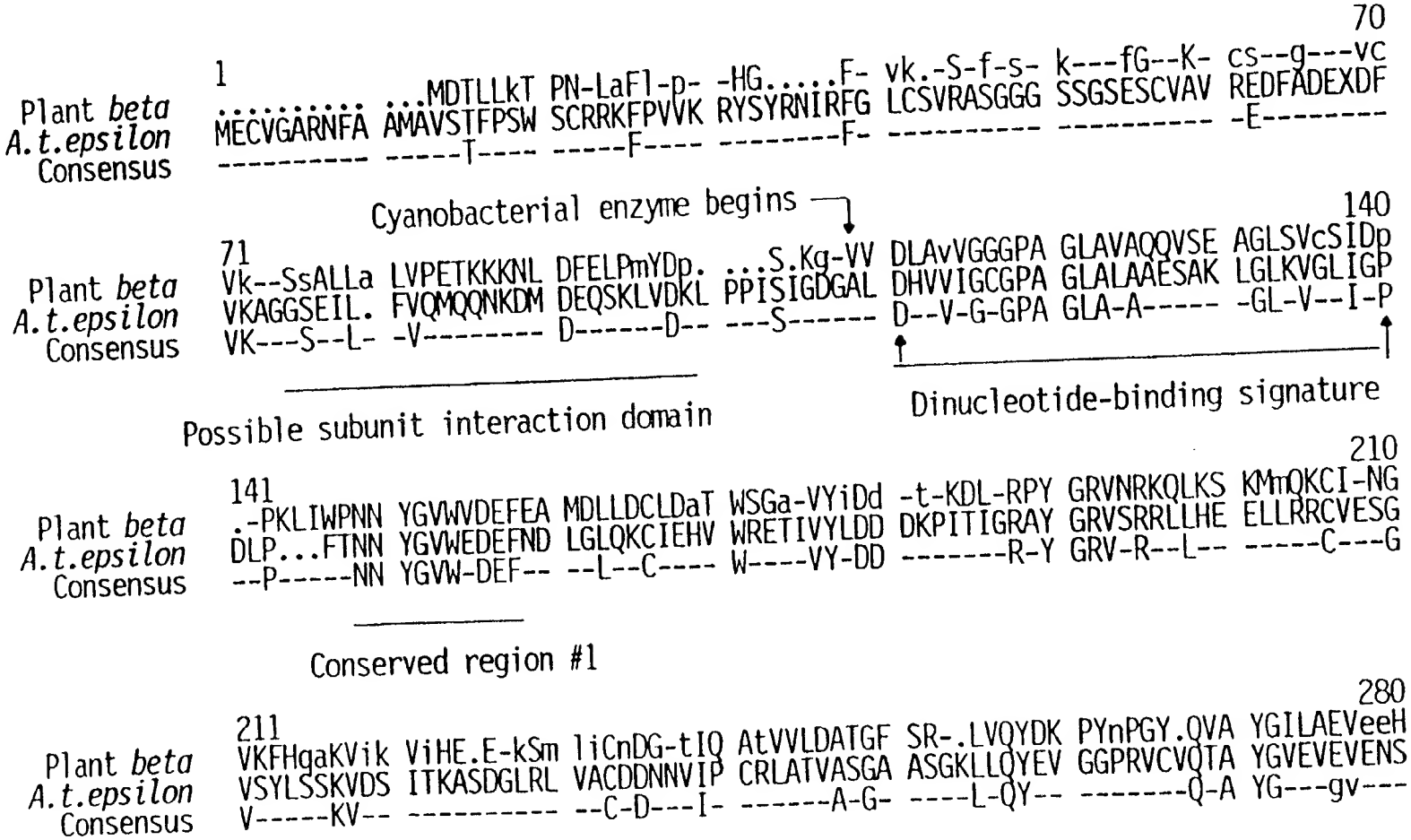
```
251                                                                300
RYVTQEELRQ MMQ....PDN GLQWSPWFRI IAARFLERWW ADLDAALNTD
RYVTQEELRQ MMQ....PDN GLQWSPWFRI IAARFLERWW ADLDAALNTD
KYVSREELKE LVKKADAGEE GLKLSPWFRL VVDNFLMKWW DHVEKGTLVE
KYMNRDDLKE LLRKADAEEE GVKLSPWFRL VVDNFLFKWW DHVEKGSLKD
KYVSREELKE LVKKADAGDE AVKLSPWFRL VVDNFLMKWW DHVEKGTITE
KWVSPNDLKT MF.....ADP SYKFTPWFKI ICENYLFNWW EQLDDLSEVE

301
KHEDWGTVHH INEA*
KHEDWGTVHH INEA*
A.IDMKTIHK L*
A.ADMKTIHK L*
A.ADMKTIHK L*
A.ADMKTIHK L*
NDRQ...IHR ML*
```

FIG. 12

```
   1  ccaaaaacaa ctcaaatctc ctccgtcgct cttactccgc catgggtgac
  51  gactccggca tggatgctgt tcagcgacgt ctcatgtttg acgatgaatg
 101  cattttggtg gatgagtgtg acaatgtggt gggacatgat accaaataca
 151  attgtcactt gatggagaag attgaaacag gtaaaatgct gcacagagca
 201  ttcagcgttt ttctattcaa ttcaaaatac gagttacttc ttcagcaacg
 251  gtctgcaacc aaggtgacat ttcctttagt atggaccaac acctgttgca
 301  gccatccact ctacagagaa tccgagcttg ttcccgaaac gcctgagaga
 351  atgctgcaca gaggaxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx
 401  xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx
 451  xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx
 501  xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx
 551  xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx
 601  xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx
 651  xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx tcatgtgcaa aagggtacac
 701  tcactgaatg caatttgata tgaaaaccat acacaagctg atatagaaac
 751  acaccctcaa ccgaaaagca agcctaataa ttcgggttgg gtcgggtcta
 801  ccatcaattg ttttttttctt ttaacaactt ttaatctcta tttgagcatg
 851  ttgattcttg tcttttgtgt gtaagatttt gggtttcgtt tcagttgtaa
 901  taatgaacca ttgatggttt gcaatttcaa gttcctatcg acatgtagtg
 951  atctaaaaaa
```

FIG. 13A


1
70
Plant beta
A.t.epsilon
Consensus
Cyanobacterial enzyme begins
71
140
Plant beta
A.t.epsilon
Consensus
Possible subunit interaction domain
Dinucleotide-binding signature
141
210
Plant beta
A.t.epsilon
Consensus
Conserved region #1
211
280
Plant beta
A.t.epsilon
Consensus

281
350
Plant beta PFD--KMVfM DWRDsHL-NN -eLKERNs-i PTFLYAMPFS SNrIFLEETS LVARPGLmd DIQERMVARL
A.t.epsilon PYDPDQMVFM DYRDY..TNE .KVRSLEAEY PTFLYAMPMT KSRLFFEXTC LASKDVMPFD LLKTKLMLRL
Consensus P-D---MVFM D-RD---N- ---------- PTFLYAMP-- --R-F-EXT- L---------D ---------RL
Conserved region #2
Conserved region #3
351
420
Plant beta -HLGIKVKsI EEDEhCvIPM GGpLPV1PQR VVGiGGTAGm VHPSTGYMVA RTLAAAPvVA NAIi-YLgSe
A.t.epsilon DTLGTRILKT YEEEWSVIPV GGSLPNTEQK NLAFGAAASM VHPATGYSVV RSLSEAPKYA SVIAEILREE
Consensus --LGI----- -E-E---IP- GG-LP---Q- ----G--A-M VHP-TGY-V- R-L--AP--A --I---L--E
Conserved region #4
Predicted TM helix
421
480
Plant beta -s-s..G-eL SaeVWkDLWP IERRRQREFF CFGMDILLKL DLpATRRFFD AFFDLaPrYW
A.t.epsilon TTKQINSN.I SRQAWDTLWP PERKRQRAFF LFGLALIVQF DTEGIRSFFR TFFRLPKMMW
Consensus ---------- S---W--LWP -ER-RQR-FF -FG-------- D----R-FF- -FF-L---W
Conserved region #5
481
533
Plant beta HGFLSSRLfL PELivFGLSL FShASNTSR- EIMTK.GT-P Lv-MINNL1Q D--
A.t.epsilon QGFLGSTLTS GDLVLFALYM FVISPNNLRK GLINHLISDP TGATMIKTYL KV.
Consensus -GFL-S-L-- --L--F-L-- F----N--R- ---------- ---------- ---
Predicted TM helix

FIG. 13B

FIG. 14A

Adonis palaestina ε-cyclase cDNA #5 Length: 1898

```
   1 aaaggagtgt tctattaatg ttactgtcgc attcttgcaa cacttatatt
  51 caaactccat tttcttcttt tctcttcaaa acaacaaact aatgtgagca
 101 gagtatctgg ctatggaact acttggtgtt cgcaacctca tctcttcttg
 151 ccctgtgtgg acttttggaa caagaaacct tagtagttca aaactagctt
 201 ataacataca tcgatatggt tcttcttgta gagtagattt tcaagtgaga
 251 gctgatggtg gaagcgggag tagaagttct gttgcttata aagagggttt
 301 tgtggatgaa gaggatttta tcaaagctgg tggttctgag cttttgtttg
 351 tccaaatgca gcaaacaaag tctatggaga aacaggccaa gctcgccgat
 401 aagttgccac caatacctt tggagaatcc gtgatggact tggttgtaat
 451 aggttgtgga cctgctggtc tttcactggc tgcagaagct gctaagctag
 501 ggttgaaagt tggccttatt ggtcctgatc ttccttttac aaataattat
 551 ggtgtgtggg aagacgagtt caaagatctt ggacttgaac gttgtatcga
 601 gcatgcttgg aaggacacca tcgtatatct tgataatgat gctcctgtcc
 651 ttattggtcg tgcatatgga cgagttagtc gacatttgct acatgaggag
 701 ttgctgaaaa ggtgtgtgga gtcaggtgta tcatatctgg attctaaagt
 751 ggaaaggatc actgaagctg gtgatggcca tagccttgta gtttgtgaaa
 801 atgagatctt tatcccttgc aggcttgcta ctgttgcatc tggagcagct
 851 tcagggaaac ttttggagta tgaagtaggt ggccctcgtg tttgtgtcca
 901 aaccgcttat ggggtggagg ttgaggtgga gaacaatcca tacgatccca
 951 acttaatggt attcatggac tacagagact atatgcaaca gaaattacag
1001 tgctcggaag aagaatatcc aacatttctC tatgtcatgc ccatgtcgcc
1051 aacaagactt ttttttgagg aaacctgttt ggcctcaaaa gatgccatgc
1101 cattcgatct actgaagaga aaactgatgt cacgattgaa gactctgggt
1151 atccaagtta caaaagttta tgaagaggaa tggtcatata ttcctgttgg
1201 tggttcttta ccaaacacag agcaaaagaa cctagcattt ggtgctgcag
1251 caagcatggt gcatccagca acaggctatt cggttgtacg gtcactgtca
1301 gaagctccaa aatatgcttc tgtaattgca aagattttga agcaagataa
1351 ctctgcgtat gtggtttctg gacaaagtag tgcagtaaac atttcaatgc
1401 aagcatggag cagtctttgg ccaaaggagc gaaaacgtca aagagcatTc
1451 tttcttttTg gattagagct tattgtgcag ctagatattg aagcaaccag
1501 aacattcttt agaaccttct tccgcttgcc aacttggatg tggtggggtt
1551 tccttgggtc ttcactatca tctttcgatc tcgtcttgtt ttccatgtac
1601 atgtttgttt tggcgccaaa cagcatgagg atgtcacttg tgagacattt
1651 gctttcagat ccttctggtg cagttatggt aagagcttac ctcgaaaggt
1701 agtctcatct attattaaac tctagtgttt caccaaataa atgaggatcc
1751 ttcgaatgtg tatatgatca tctctatgta tatcctgtac tctaatctca
1801 taaagtaaat gccgggtttg atattgttgt gtcaaaccgg ccaatgatat
1851 aaagtaaatt tattgataca aaagtagttt ttttccttaa aaaaaaaa
```

FIG. 14B

Adonis palaestina ε-cyclase #5 predicted polypeptide
TRANSLATE from: 113 to: 1702 Length: 529 amino acids

```
  1  MELLGVRNLI SSCPVWTFGT RNLSSSKLAY NIHRYGSSCR VDFQVRADGG
 51  SGSRSSVAYK EGFVDEEDFI KAGGSELLFV QMQQTKSMEK QAKLADKLPP
101  IPFGESVMDL VVIGCGPAGL SLAAEAAKLG LKVGLIGPDL PFTNNYGVWE
151  DEFKDLGLER CIEHAWKDTI VYLDNDAPVL IGRAYGRVSR HLLHEELLKR
201  CVESGVSYLD SKVERITEAG DGHSLVVCEN EIFIPCRLAT VASGAASGKL
251  LEYEVGGPRV CVQTAYGVEV EVENNPYDPN LMVFMDYRDY MQQKLQCSEE
301  EYPTFLYVMP MSPTRLFFEE TCLASKDAMP FDLLKRKLMS RLKTLGIQVT
351  KVYEEEWSYI PVGGSLPNTE QKNLAFGAAA SMVHPATGYS VVRSLSEAPK
401  YASVIAKILK QDNSAYVVSG QSSAVNISMQ AWSSLWPKER KRQRAFFLFG
451  LELIVQLDIE ATRTFFRTFF RLPTWMWWGF LGSSLSSFDL VLFSMYMFVL
501  APNSMRMSLV RHLLSDPSGA VMVRAYLER*
```

FIG. 15A

DNA sequence of potato cDNA (GenBank R27545) obtained from Nicholas J. Provart

```
potato.seq  Length: 1378  August 2, 1996  13:06  Type: N  Check: 605  ..
   1  tagcggnnnn naggatgagt tcaaagatct tggtcttcaa gcctgcattg
  51  aacatgtttg gcgggatacc attgtatatc ttgatgatga tgatcctatt
 101  cttattggcc gtgcctatgg aagagttagt cgccatttac tgcacgagga
 151  gttactcaaa aggtgtgtgg aggcaggtgt tttgtatcta aactcgaaag
 201  tggataggat tgttgaggcc acaaatggcc acagtcttgt agagtgcgag
 251  ggtgatgttg tgattccctg caggtttgtg actgttgcat cgggagcagc
 301  ctcggggaaa ttcttgcagt atgagttggg aggtcctaga gtttctgttc
 351  aaacagctta tggagtggaa gttgaggtcg ataacaatcc atttgacccg
 401  agcctgatgg ttttcatgga ttatagagac tatgtcagac acgacgctca
 451  atctttagaa gctaaatatc caacatttct ctatgccatg cccatgtctc
 501  caacacgagt ctttttcgag gaaacttgtt tggcttcaaa agatgcaatg
 551  ccattcgatc tgttaaagaa aaaattgatg ttacgattga acaccctcgg
 601  tgtaagaatt aaagaaattt atgaggagga atggtcttac ataccagttg
 651  gaggatcttt gccaaataca gaacaaaaaa cacttgcatt tggtgctgct
 701  gctagcatgg ttcatccagc cacaggttat tcagtcgtca gatcactgtc
 751  tgaagctcca aaatgcgcct tcgtgcttgc aaatatatta cgacaaaatc
 801  atagcaagaa tatgcttact agttcaagta ccccgagtat ttcaactcaa
 851  gcttggaaca ctctttggcc acaagaacga aaacgacaaa gatcgttttt
 901  cctatttgga ctggctctga tattgcagct ggatattgag gggataaggt
 951  cattttccg cgcgttcttc cgtgtgccaa aatggatgtg gcagggattt
1001  cttggttcaa gtctttcttn agcagacctc atgttatttg ccttctacat
1051  gtttattatt gcaccaaatg acatgagaag aggcttaatc agacatcttt
1101  tatctgatcc tactggtgca acattgataa gaacttatct tacattttag
1151  agtaaattcc tcctacaata gttgttgaan nagaggcctc attacttcag
1201  attcataaca gaaatcgcgg tctctcgagg ccttgtatat aacattttca
1251  ctaggttaat attgcttgaa taagttgcac agtttcagtt tttgtatctg
1301  cttctttttt gtccaagatc atgtattgan ccaatttata tacattgcca
1351  gtatatataa attttataaa aaaaaaaa

poteps.pep  Length: 378   TRANSLATE from: 14  to: 1147
   1  DEFKDLGLQA CIEHVWRDTI VYLDDDDPIL IGRAYGRVSR HLLHEELLKR
  51  CVEAGVLYLN SKVDRIVEAT NGHSLVECEG DVVIPCRFVT VASGAASGKF
 101  LQYELGGPRV SVQTAYGVEV EVDNNPFDPS LMVFMDYRDY VRHDAQSLEA
 151  KYPTFLYAMP MSPTRVFFEE TCLASKDAMP FDLLKKKLML RLNTLGVRIK
 201  EIYEEEWSYI PYGGSLPNTE QKTLAFGAAA SMVHPATGYS VVRSLSEAPK
 251  CAFVLANILR QNHSKNMLTS SSTPSISTQA WNTLWPQERK RQRSFFLFGL
 301  ALILQLDIEG IRSFFRAFFR VPKWMWQGFL GSSLSXADLM LFAFYMFIIA
 351  PNDMRRGLIR HLLSDPTGAT LIRTYLTF*
```

FIG. 15B

Chimeric lettuce/potato lycopene ε-cyclase: converts lycopene to δ-carotene, the lettuce cDNA converts lycopene to ε-carotene and the potato cDNA does not produce an active enzyme

(amino acids in lower case are from lettuce and those in uppercase are from the potato cDNA; an *Ava*II site in common to the two cDNAs was used to construct the chimera)

```
  1  mecfgarnmt atmavftcpt ftdcnirhkf sllkqrrftn lsassslrqi
 51  kcsaksdrcv vdkqgisvac eedyvkaggs elffvqmqrt ksmesqsklS
101  eklaqipign cildlvvigc gpaglalaae saklglnvgl igpdlpftnn
151  ygvwqdefig lglegciehs wkdtlvyldd adpirigray grvhrdllhe
201  ellrrcvesg vsylsskver iteapngysl iecegnitip crlatvasga
251  asgkfleyel gGPRVSVQTA YGVEVEVDNN PFDPSLMVFM DYRDYVRHDA
301  QSLEAKYPTF LYAMPMSPTR VFFEETCLAS KDAMPFDLLK KKLMLRLNTL
351  GVRIKEIYEE EWSYIPVGGS LPNTEQKTLA FGAAAASMVHP ATGYSVVRSL
401  SEAPKCAFVL ANILRQNHSK NMLTSSSTPS ISTQAWNTLW PQERKRQRSF
451  FLFGLALILQ LDIEGIRSFF RAFFRVPKWM WQGFLGSSLS XADLMLFAFY
501  MFIIAPNDMR RGLIRHLLSD PTGATLIRTY LTF*
```

09/701395

FIG. 16

GAP comparison of Arabidopsis ε-cyclase x potato ε-cyclase (partial)
blosum62.cmp Gap Weight: 12 Average Match: 2.912
Length Weight: 4 Average Mismatch: -2.003
Quality: 1485 Length: 529
Ratio: 3.929 Gaps: 1
Percent Similarity: 79.893 Percent Identity: 76.139
Match display thresholds for the alignment(s):
| = IDENTITY : = 2 . = 1

```
151  EDEFNDLGLQKCIEHVWRETIVYLDDDKPITIGRAYGRVSRRLLHEELLR 200
     |||  |||||  |||||||:||||||||  ||  ||||||||||  |||||||:
  1  .DEFKDLGLQACIEHVWRDTIVYLDDDDPILIGRAYGRVSRHLLHEELLK  49

201  RCVESGVSYLSSKVDSITEASDGLRLVACDDNNVIPCRLATVASGAASGK 250
     ||||.||  ||.||||  |  ||..|  ||  |: .  |||||  ||||||||||
 50  RCVEAGVLYLNSKVDRIVEATNGHSLVECEGDVVIPCRFVTVASGAASGK  99

251  LLQYEVGGPRVCVQTAYGVEVEVENSPYDPDQMVFMDYRDYTNEKVRSLE 300
     ||||.|||||  ||||||||||:|.|:||  |||||||||      .|||
100  FLQYELGGPRVSVQTAYGVEVEVDNNPFDPSLMVEMDYRDYVRHDAQSLE 149

301  AEYPTFLYAMPMTKSRLFFEETCLASKDVMPFDLLKTKLMLRLDTLGIRI 350
     |.||||||||||. .|.|||||||||| |||||||  ||||||.|||:||
150  AKYPTFLYAMPMSPTRVFFEETCLASKDAMPFDLLKKKLMLRLNTLGVRI 199

351  LKTYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAP 400
     . ||||||||||||||||||| |||||||||||||||||||||||||||||
200  KEIYEEEWSYIPVGGSLPNTEQKTLAFGAAASMVHPATGYSVVRSLSEAP 249

401  KYASVIAEILREETTKQI.....NSNISRQAWDTLWPPERKRQRAFFLFG 445
     | | |:| |||:  .| .       .|| |||.|||| ||||||.|||||
250  KCAFVLANILRQNHSKNMLTSSSTPSISTQAWNTLWPQERKRQRSFFLFG 299

446  LALIVQFDTEGIRSFFRTFFRLPKWMWQGFLGSTLTSGDLVLFALYMFVI 495
     ||||.|  | ||||||||  |||.||||||||||||.|.  ||.|||  |||:|
300  LALILQLDIEGIRSFFRAFFRVPKWMWQGFLGSSLSXADLMLFAFYMFII 349

496  SPNNLRKGLINHLISDPTGATMIKTYLKV 524
     .||.:|:|||  ||:|||||||:|:|||
350  APNDMRRGLIRHLLSDPTGATLIRTYLTF 378
```

09/701395

FIG. 17A

Adonis palaestina Ipi1

```
   1 attcatcttc agcagcgctg tcgtactctt tctatatctt cttccatcac
  51 taacagtagt cgccgacggt tgaatcggct attcgcctca acgtcaacta
 101 tgggtgaagt cactgatgct ggaatggatg ctgttcagaa gcggctcatg
 151 ttcgacgacg aatgtatttt ggtggatgag aatgacaagg tcgtcgggca
 201 tgattccaaa tacaactgtc atttgatgga aaagatagag gcagaaaatt
 251 tgcttcacag agccttcagt gttttcttgt tcaactcaaa atatgaattg
 301 cttcttcagc aacgatccgc cacaaaggta acattcccgc tcgtatggac
 351 aaacacatgt tgcagtcatc ctctctttcg tgattccgag ctcatagaag
 401 aaaattatct cggtgtacga aacgctgcac aaagaaagct tttagacgag
 451 ctaggcattc cagctgaaga tgtcccagtt gatgaattta ctcctcttgg
 501 tcgcattctt tacaaagctc catctgacgg caaatgggga gagcacgaat
 551 tggactatct cctatttatt gtccgagatg tgaaatacga tccaaaccca
 601 gatgaagttg ctgatgctaa gtatgttaat cgcgaggagt tgagagagat
 651 actgagaaaa gctgatgctg gtgaagaggg actcaagttg tctccttggt
 701 ttagattggt tgttgataac tttttgttca agtggtggga tcatgtagag
 751 cagggtacga ttaaggaagt tgctgacatg aaaactatcc acaagttgac
 801 ttaagaggac ttctctcctc tgttctacta tttgtttttt gctacaataa
 851 gtgggtggtg ataagcagtt tttctgtttt ctttaattta tggcttttga
 901 atttgcctcg atgttgaact tgtaacatat ttagacaaat atgagacctt
 951 gtaagttgaa tttgaggctg aatttatatt tttgggaaca taataatgtt
1001 aa
```

FIG. 17B

Adonis palaestina Ipi2

```
   1  ttttaaagct ctttcgctcc accaccatca aagccagcca aatttctctg
  51  tacaaaagtt aaaaacaccg ctttgggctt tggcccctcc atatcggaat
 101  ccttgtttac gatacgcatc taaaccagta attctcggtt ttaatttgtt
 151  tcctaaatta ggcccctttc cggaatcccg agaattatgt cgtcgatcag
 201  gattaatcct ttatatagta tcttctccac caccactaaa acattatcag
 251  cttcgtgttc ttctcccgct gttcatcttc agcagcgttg tcgtactctt
 301  tctatttctt cttccatcac taacagtcct cgccgagggt tgaatcggct
 351  gttcgcctca acgtcgacta tgggtgaagt cgctgatgct ggtatggatg
 401  ccgtccagaa gcggcttatg ttcgacgatg aatgtatttt ggtggatgag
 451  aatgacaagg tcgtcggaca tgattccaaa tacaactgtc atttgatgga
 501  aaagatagag gcagaaaact tgcttcacag agccttcagt gttttcttat
 551  tcaactcaaa atacgagttg cttcttcagc aacgatctgc aacgaaggta
 601  acattcccgc tcgtatggac aaacacctgt tgcagccatc ccctcttccg
 651  tgattccgaa ctcatagaag aaaattttct cggggtacga aacgctgcac
 701  aaaggaagct tttagacgag ctaggcattc cagctgaaga cgtaccagtt
 751  gatgaattca ctcctcttgg tcgcattctt tacaaagctc catctgacgg
 801  aaaatgggga gagcacgaac tggactatct tctgtttatt gtccgagatg
 851  tgaaatacga tccaaaccca gatgaagttg ctgacgctaa gtacgttaat
 901  cgcgaggagt tgaaagagat actgagaaaa gctgatgcag gtgaagaggg
 951  aataaagttg tctccttggt ttagattggt tgtggataac tttttgttca
1001  agtggtggga tcatgtagag gaggggaaga ttaaggacgt cgccgacatg
1051  aaaactatcc acaagttgac ttaagagaaa gtctcttaag ttctactatt
1101  tggtttttgc ttcaataagt ggatggtgat gagcagtttt tatgcttcct
1151  ttaattttgg cttttcaatt tgctttatgt gttgaacttg taacatattt
1201  agtcaaatat gagaccttgt gagttgaatt tgaggttata tttatagttt
1251  tgggaacata aaaaaaaaaa
```

FIG. 18A

Haematococcus pluvialis Ipil

```
   1  ctcggtagct ggccacaatc gctatttgga acctggcccg gcggcagtcc
  51  gatgccgcga tgcttcgttc gttgctcaga ggcctcacgc atatcccccg
 101  cgtgaactcc gcccagcagc ccagctgtgc acacgcgcga ctccagttta
 151  agctcaggag catgcagatg acgctcatgc agcccagcat ctcagccaat
 201  ctgtcgcgcg ccgaggaccg cacagaccac atgaggggtg caagcacctg
 251  ggcaggcggg cagtcgcagg atgagctgat gctgaaggac gagtgcatct
 301  tggtggatgt tgaggacaac atcacaggcc atgccagcaa gctggagtgt
 351  cacaagttcc taccacatca gcctgcaggc ctgctgcacc gggccttctc
 401  tgtgttcctg tttgacgatc aggggcgact gctgctgcaa cagcgtgcac
 451  gctcaaaaat caccttccca agtgtgtgga cgaacacctg ctgcagccac
 501  cctttacatg ggcagacccc agatgaggtg gaccaactaa gccaggtggc
 551  cgacggaaca gtacctggcg caaaggctgc tgccatccgc aagttggagc
 601  acgagctggg gataccagcg caccagctgc cggcaagcgc gtttcgcttc
 651  ctcacgcgtt tgcactactg tgccgcggac gtgcagccag ctgcgacaca
 701  atcagcgctc tggggcgagc acgaaatgga ctacatcttg ttcatccggg
 751  ccaacgtcac cttggcgccc aaccctgacg aggtggacga agtcaggtac
 801  gtgacgcaag aggagctgcg gcagatgatg cagccggaca acgggctgca
 851  atggtcgccg tggtttcgca tcatcgccgc gcgcttcctt gagcgttggt
 901  gggctgacct ggacgcggcc ctaaacactg acaaacacga ggattgggga
 951  acggtgcatc acatcaacga agcgtgaaag cagaagctgc aggatgtgaa
1001  gacacgtcat ggggtggaat tgcgtacttg gcagcttcgt atctcctttt
1051  tctgagactg aacctgcagt caggtcccac aaggtcaggt aaaatggctc
1101  gataaaatgt accgtcactt tttgtcgcgt atactgaact ccaagaggtc
1151  aaaaaaaaaa aaaaa
```

09/701395

FIG. 18B

Haematococcus pluvialis Ipi2

```
   1 tggaacctgg cccggcggca gtccgatgcc gcgatgcttc gttcgttgct
  51 cagaggcctc acgcatatcc cgcgcgtgaa ctccgcccag cagcccagct
 101 gtgcacacgc gcgactccag tttaagctca ggagcatgca gctgcttgcc
 151 gaggaccgca cagaccacat gaggggtgca agcacctggg caggcgggca
 201 gtcgcaggat gagctgatgc tgaaggacga gtgcatctta gtggatgctg
 251 acgacaacat cacaggccat gccagcaagc tggagtgcca caaattccta
 301 ccacatcagc ctgcaggcct gctgcaccgg gccttctctg tgttcctgtt
 351 tgacgaccag gggcgactgc tgctgcaaca gcgtgcacgc tcaaaaatca
 401 ccttcccaag tgtgtggacg aacacctgct gcagccaccc tctacatggg
 451 cagaccccag atgaggtgga ccaactaagc caggtggccg acggcacagt
 501 acctggcgca aaagctgctg ccatccgcaa gttggagcac gagctgggga
 551 taccagcgca ccagctgccg gcaagcgcgt ttcgcttcct cacgcgtttg
 601 cactactgtg ccgcggacgt gcagccggct gcgacacaat cagcgctctg
 651 gggcgagcac gagatggact acatcttatt catccgggcc aacgtcacct
 701 tggcgcccaa ccctgacgag gtggacgaag tcaggtacgt gacgcaagag
 751 gagctgcggc agatgatgca gccggacaac gggttgcaat ggtcgccgtg
 801 gtttcgcatc atcgccgcgc gcttccttga gcgttggtgg gctgacctgg
 851 acgcggccct aaacactgac aaacacgagg attggggaac ggtgcatcac
 901 atcaacgaag cgtgaaggca gaagctgcag gatgtgaaga cacgtcatgg
 951 ggtggaattg cgtacttggc agcttcgtat ctcctttttc tgagactgaa
1001 cctgcagagc tagagtcaat ggtgcatcat attcatcgtc tctcttttgt
1051 tttagactaa tctgtagcta gagtcactga tgaatccttt acaactttca
1101 aaaaaaaaa
```

FIG. 19A

Lactuca sativa Ipi1

```
   1  tgccaaaatg ttgaaatttc ccccttttaa aaccattgct accatgatct
  51  cttctccata ttcttccttc ttgctgcctc ggaaatcttc tttccctcca
 101  atgccgtctc tcgcagccgc tagtgttttc ctccaccctc tttcgtctgc
 151  cgctatgggc gattccagca tggatgctgt ccagcgacgt ctcatgttcg
 201  atgacgaatg cattttggtg gatgagaatg acaaagtggt tggccatgat
 251  actaaataca attgtcattt gatggagaag attgaaaagg gaaatatgct
 301  acacagagca ttcagtgtgt tcttgttcaa ctcgaaatat gaattactcc
 351  ttcagcaacg ttctgcaacc aaggtgactt tccctttggt atggacaaac
 401  acgtgttgca gccatccact atacagggag agtgagctta ttgacgaaaa
 451  cgcccttggg gtgaggaatg ctgcacagag gaagctcctg gatgaactcg
 501  gcatccctgg agcagatgtt ccggttgatg agttcactcc attgggtcgc
 551  attctataca aggccgcatc ggatggaaag tggggagaac atgaacttga
 601  ttacctgctg tttatggtac gtgatgttgg tttggatccg aacccagatg
 651  aagtgaaaga tgtaaaatat gtgaaccggg aagagctgaa ggaattggta
 701  aggaaggcgg atgctggtga agagggtgtg aagctgtccc cgtggttcaa
 751  attgattgtc gataatttct tgtttcagtg gtgggatcga ctccataagg
 801  gaaccctaac cgaagctatt gatatgaaaa caatccacaa actcacataa
 851  aaacactaca ctagtaggag agaggattat atgagatatt tgttatatgt
 901  gaaattgaaa ttcagatgaa tgcttgtatt tatttctatt tggacaaact
 951  tcaacttctt tttgctacct tatcagaaaa aaaaa
```

FIG. 19B

Lactuca sativa Ipi2

```
   1  tattcgcttc aaaatctctt ccattaactg ctcaaatctc caccttcgcc
  51  ggtcttaatc tccgccggcg cactttcacc accataaccg ccgccatggg
 101  tgacgattcc ggcatggacg ctgtccagag acgtctcatg tttgatgatg
 151  aatgcatttt ggttgatgaa aatgacaatg ttcttgggca tgataccaaa
 201  tacaattgtc acttgatgga gaagattgag aaagataatt tgcttcatag
 251  agcattcagt gtatttttat tcaattcaaa atacgaatta ctccttcagc
 301  aaaggtcaga aaccaaggtg acatttcctt tggtatggac aaacacctgt
 351  tgcagccatc cactatacag agaatcggag ttaattcccg aaaatgccct
 401  tggggtcaga aatgctgcac agaggaagct tctagatgaa ctcggtatcc
 451  ctgctgaaga tgttccagtt gatgagttca caactttagg tcgcatgttg
 501  tacaaggctc catctgatgg aaaatggggt gaacatgaag ttgattacct
 551  actcttcctc gtgcgtgacg ttgccgtgaa cccaaaccct gatgaggtgg
 601  cggacattag atacgtgaac caagaagagt taaaagagtt actaaggaag
 651  gcggatgcgg gtgaggaggg tttgaaattg tccccatggt ttaggctagt
 701  ggtggacaac ttcttgttca aatggtggga tcatgtccaa aaggggacac
 751  tcaatgaagc aattgacatg aaaaccattc ataagttgat atgaaaaatg
 801  gttaatattt atggtggtgg tttggagcta ataatttgtg tgttcaagtc
 851  tcggtccttc ttttttaac gttttttttt tttcttttat tgggagtgtt
 901  tattgtgtac ttgtaacgta ggccctttgg ttacgcttta agagtttaat
 951  aaagaaccac cgttaattta aaaaaaaaaa aaaaaaaa
```

FIG. 20

Chlamydomonas reinhardtii Ipi1

(Note: the isomerase cDNA probably ends at ca. base 1103; the second half of the cDNA is similar to extensin and other hydroxyproline-rich structural proteins)

```
   1  ggcacgagct cgagtttgtt ttaccatgac atcgggaatt tggaagcttg
  51  aactacctca attactcaag taactcgcgg caacacattt cgcgcgccat
 101  cgctgttttc tctgctccag ctaccgagca gcattgcttt agatcgcttt
 151  gatgtcataa actcccactt atatgagatc cagtttcatc gagcccaagc
 201  ccagagcgca acctgtctta agccgcggca gggcgtccat gcgcctcgcg
 251  caaagccgtg ctctcgttgc gcgtgtcagc tccgccctgt ggccgggagc
 301  aggactttca caggctcaaa gcgttgcggt gcgaatggcg agttcgtcaa
 351  cctgggaagg cacgggcctg agccaggatg acttcatgca gcgggacgag
 401  tgcttggtgg tggacgagca ggaccggctg ctaggcaccg ccaacaagta
 451  cgactgccac cgcttcgagg cggccaaggg ccagccctgc ggccgcctgc
 501  accgcgcctt ctccgtgttc ctgttcagcc ccgacggccg actgctgctg
 551  cagcagcgcg cagccagcaa ggtgacgttc ccgggtgtgt ggaccaacac
 601  ctgctgctcg cacccgctgg cgggccaggc gccggacgag gtggacctgc
 651  cggcggcggt agcctcgggc caggtgccgg gcatcaaggc ggcggcggtg
 701  cgcaagctgc agcacgagct ggggataccg ccggagcagg ttcccgcctc
 751  ctccttctcc ttcctcacgc gtctgcacta ctgcgccgcc gacaccgcca
 801  cgcacggccc ggcggcggag tggggcgagc acgaggtgga ctacgtgctg
 851  ttcgtgcggc cgcagcagcc cgtcagcctg cagcccaacc cagacgaggt
 901  ggacgccacg cgctacgtga cgctgccgga gcttcagtcc atgatggcgg
 951  accccggcct cagctggagc ccctggttcc gcatcctggc cacacagccc
1001  gccttcctgc ccgcctggtg gggcgacctg aagcggcgct ggcgcccggg
1051  cggcagccga ctgtaggact ggggcaccat ccaccgcgtc atgtgaagaa
1101  aaaggggaag caggggcggg agcgggggat gaatgggaat gtgaatgcga
1151  ttgtgatgcg gcgtgggatg aggtctgaag acaggggggaa aatcgggggg
1201  cgggcgtgag cgtgtgtgta cgtgagcgac aaagccggga ggcggaccgc
1251  gcgatgggta catgtgtgtg cggagggtcg gtgggtcggt cggttgcgcg
1301  gcatagcgtg ttgtgtgtgt gcggctgcgc gggtatgtgg gcacccgggc
1351  acggaggaga aggcacacgc aggtggcgcg gaggtgtgtc aggggccatg
1401  ggcgggcctc actcctggtc gtgcccagtg gtctcgtggg cagagtggca
1451  ggggctgcac ccatatgagc ggcgcactgc cgcgctgggc taagtcctta
1501  tcacttggtg aggtggggcg aggtggctgt gggcggcggg cgcagtggca
1551  gaaggacacg gtgtgtgagc ggtggagctc tggccgtgcc ggccgtgagg
1601  ggcggatagc gatatgacgt tgtgcttggc cgctgtaatg cgggagaatg
1651  tgcaggccgc gagaagcggg cggtggcagg aggccgcagg ctgcagcacc
1701  cgttggggag gtgccgcctg caggcgcggc gccgggcggg cctgagtaat
1751  gggcgcctga gtagtggcgg ccacaggagg cgcaggaggc agcagcagga
1801  ggacgagctg gagggacccg ttggcaaccc aaggttgcgc gtgtaacata
1851  gtggccatac aaaaaaaaaa aaaa
```

FIG. 21A

Tagetes erecta Ipi1

```
   1  ccaaaaacaa ctcaaatctc ctccgtcgct cttactccgc catgggtgac
  51  gactccggca tggatgctgt tcagcgacgt ctcatgtttg acgatgaatg
 101  cattttggtg gatgagtgtg acaatgtggt gggacatgat accaaataca
 151  attgtcactt gatggagaag attgaaacag gtaaaatgct gcacagagca
 201  ttcagcgttt ttctattcaa ttcaaaatac gagttacttc ttcagcaacg
 251  gtctgcaacc aaggtgacat ttcctttagt atggaccaac acctgttgca
 301  gccatccact ctacagagaa tccgagcttg ttcccgaaaa cgcccttgga
 351  gtaagaaatg ctgcacagag gaagctgttg gatgaactcg gtatccctgc
 401  tgaagatgtt cccgttgatc agtttactcc tttaggtcgc atgctctaca
 451  aggctccatc tgatggaaag tggggagaac atgaacttga ctacctactt
 501  ttcatagtga gagacgttgc tgtaaacccg aacccagatg aagtggcgga
 551  tatcaaatat gtganccang aagagttaaa ggagctgcta aggaaagcag
 601  atgcggggga ggagggtttg aagctgtctc catggttcag gttagtggtt
 651  gataacttct tgttcaagtg gtgggatcat gtgcaaaagg gtacactcac
 701  tgaagcaatt gatatgaaaa ccatacacaa gctgatatag aaacacaccc
 751  tcaaccgaaa agttcaagcc taataattcg ggttgggtcg ggtctaccat
 801  caattgtttt tttcttttaa gaagttttaa tctctatttg agcatgttga
 851  ttcttgtctt ttgtgtgtaa gattttgggt ttcgtttcag ttgtaataat
 901  gaaccattga tggtttgcaa tttcaagttc ctatcgacat gtagtgatct
 951  aaaaaa
```

FIG. 21B

Oryza sative Ipi1

```
   1  cctccctttg cctcgcgcag aggcggccgc gccttctccg ccgcgaggat
  51  ggccggcgcc gccgccgccg tggaggacgc cgggatggac gaggtccaga
 101  agcggctcat gttcgacgac gaatgcattt tggtggatga acaagacaat
 151  gttgttggcc atgaatcaaa atataactgc catctgatgg aaaaaatcga
 201  atctgaaaat ctacttcata gggctttcag tgtattcctg ttcaactcaa
 251  aatatgaact cctactccag caacgatctg caacaaaggt tacatttcct
 301  ctagtttgga ccaacacttg ctgcagccat cctctgtacc gtgagtctga
 351  gcttatacag gaaaactacc ttggtgttag aaatgctgct cagaggaagc
 401  tcttggatga gctgggcatc ccagctgaag atgtgccagt tgaccaattc
 451  acccctcttg gtcggatgct ttacaaggcc ccatctgatg gaaaatgggg
 501  tgaacacgag cttgactacc tgctgttcat cgtccgcgac gtgaaggtag
 551  tcccgaaccc ggacgaagtg gccgatgtga aatacgtgag ccgtgagcag
 601  ctgaaggagc tcatccgcaa agcggacgcc ggagaggaag gcctgaagct
 651  gtctccctgg ttccggctgg ttgttgacaa cttcctcatg ggctggtggg
 701  atcacgtcga gaaaggcacc ctcaacgagg ccgtggacat ggagaccatc
 751  cacaagctga agtaaggact gcgatgttgt ggctggaaag aatgatcctg
 801  aagactctgt tcttgtgctg ctgcatatta ctcttaccag ggaagttgca
 851  gaagtcagaa gaagcttttg tatgtttctg ggtttggagc ttggaagtgt
 901  tgggctctgc tgactgagag attcccttat agagtgtcta tgttaattta
 951  gcaaacttct atattataca tgattagtta attgttcggt gtctgaataa
1001  agaacaatag catgttccat gtttatttgc t
```

ClustalW 1.7 Multiple Sequence Alignment of Plant and Green Algal Isopentenyl Pyrophosphate Isomerases (IPI)
These amino acid sequences were predicted by cDNAs that were isolated and identified by color complemtation in *E.coli*

```
                    1               15 16              30 31              45 46              60 61              75 76              90
1 T.erecta 1        ---------------- ---------------- ---------------- ---------------- -MGD--DSGMDAVQR RLMFDDECILVDECD    27
2 L.sativa 1        --------MLKFPPF KTIATMISSPYSSFL LPRKSSFPPMPS--- -LAAASVFLHPLSSA AMGD--SS-MDAVQR RLMFDDECILVDEND    75
3 L.sativa 2        ---------------- ---------------- ---------------- ---------------- -MGD--DSGMDAVQR RLMFDDECILVDEND    27
4 A.palaestina 2    MSSIRINPLYSIFST TTKTLSASCSSPAVH LQQRCRTLSISSSIT NSPRRGLNRLFASTS TMGEVADAGMDAVQK RLMFDDECILVDEND    90
5 A.palaestina 1    ---------------- ---------------- ---------------- ---------------- -MGEVTDAGMDAVQK RLMFDDECILVDEND    29
6 O.sativa 1        ---------------- ---------------- ---------------- ----------MAG AAAAVEDAGMDEVQK RLMFDDECILVDEQD    33
7 A.thaliana 1      ---MSVSSLFNLPLI RLRSLALSSSFSSFR FAHRP--LSSIS--- --PRKLPNFRAFSGT AMTDTKDAGMDAVQR RLMFEDECILVDETD    80
8 A.thaliana 2      ---------------- ---------------- ---------------- ---------------- -MTDSNDAGMDAVQR RLMFEDECILVDENN    29
9 H.pluvialis 1     ---MLRSLLRGLTHI PRVNSAQQPSCAHAR LQFKLRSMQLLS--- ---------EDRTDH MRGASTWAG-GQSQD ELMLKDECILVDVED    74
10 H.pluvialis 2    ---MLRSLLRGLTHI PRVNSAQQPSCAHAR LQFKLRSMQMTLMQP SISANLSRAEDRTDH MRGASTWAG-GQSQD ELMLKDECILVDVED    86
11 C.reinhardtii 1  ---MRSSFIE---PK PRAQPVLSRGRASMR LAQSRALVARVSSAL WPGAGLSQAQSVAVR MASSSTWEGTGLSQD DFMQRDECLVVDEQD    84

                    91              105 106             120 121             135 136             150 151             165 166             180
1 T.erecta 1        NVVGHDTKYNCHLME KIE--TGKMLHRAFS VFLENSKYELLLQQR SATKVTFPLVWTNTC CSHPLYRES------ --ELVPENALGVRNA    107
2 L.sativa 1        KVVGHDTKYNCHLME KIE--KGNMLHRAFS VFLENSKYELLLQQR SATKVTFPLVWTNTC CSHPLYRES------ --ELIDENALGVRNA    155
3 L.sativa 2        NVLGHDTKYNCHLME KIE--KDNLLHRAFS VFLFNSKYELLLQQR SETKVTFPLVWTNTC CSHPLYRES------ --ELIPENALGVRNA    107
4 A.palaestina 2    KVVGYDSKYNCHLME KIE--AENLLHRAFS VFLFNSKYELLLQQR SATKVTFPLVWTNTC CSHPLFRDS------ --ELIEENFLGVRNA    170
5 A.palaestina 1    KVVGHDSKYNCHLME KIE--AENLLHRAFS VFLFNSKYELLLQQR SATKVTFPLVWTNTC CSHPLFRDS------ --ELIEENYLGVRNA    109
6 O.sativa 1        NVVGHESKYNCHLME KIE--SENLLHRAFS VFLFNSKYELLLQQR SATKVTFPLVWTNTC CSHPLYRES------ --ELIQENYLGVRNA    113
7 A.thaliana 1      RVVGHVSKYNCHLME NIE--AKNLLHRAFS VFLFNSKYELLLQQR SNTKVTFPLVWTNTC CSHPLYRES------ --ELIQDNALGVRNA    160
8 A.thaliana 2      RVVGHDTKYNCHLME KIE--AENLLHRAFS VFLFNSKYELLLQQR SKTKVTFPLVWTNTC CSHPLYRES------ --ELIEENVLGVRNA    109
9 H.pluvialis 1     NITGNASKLECHKFL PH--QPAGLLHRAFS VFLFDDQGRLLLQQR ARSKITFPSVWTNTC CSHPLHGQTPDEVDQ LSQVADGTVPGAKAA    162
10 H.pluvialis 2    NITGHASKLECHKFL PH--QPAGLLHRAFS VFLFDDQGRLLLQQR ARSKITFPSVWTNTC CSHPLHGQTPDEVDQ LSQVADGTVPGAKAA    174
11 C.reinhardtii 1  RLLGTANKYDCHRFE AAKGQPCGRLHRAFS VFLFSPDGRLLLQQR AASKVTFPGVWTNTC CSHPLAGQAPDEVDL PAAVASGQVPGIKAA    174
```

FIG. 22A

```
                    181        195 196        210 211        225 226        240 241        255 256        270
 1 T.erecta 1       AQRKLLDELGIPAED VPVDQFTPLGRMLY- --KAPSDG----KWG EHELDYLLFIVRD-- VAVNPNPDEVADIKY VSHEELKELLRKADA   188
 2 L.sativa 1       AQRKLLDELGIPGAD VPVDEFTPLGRILY- --KAASDG----KWG EHELDYLLFMVRD-- VGLDPNPDEVKDVKY VNREELKELVRKADA   236
 3 L.sativa 2       AQRKLLDELGIPAED VPVDEFTTLGRMLY- --KAPSDG----KWG EHEVDYLLFLVRD-- VAVNPNPDEVADIRY VNQEELKELLRKADA   188
 4 A.palaestina 2   AQRKLLDELGIPAED VPVDEFTPLGRILY- --KAPSDG----KWG EHELDYLLFIVRD-- VKYDPNPDEVADAKY VNREELKEILRKADA   251
 5 A.palaestina 1   AQRKLLDELGIPAED VPVDEFTPLGRILY- --KAPSDG----KWG EHELDYLLFIVRD-- VKYDPNPDEVADAKY VNREELREILRKADA   190
 6 O.sativa 1       AQRKLLDELGIPAED VPVDQFTPLGRMLY- --KAPSDG----KWG EHELDYLLFIVRD-- VKVVPNPDEVADVKY VSREQLKELIRKADA   194
 7 A.thaliana 1     AQRKLLDELGIVAED VPVDEFTPLGRMLY- --KAPSDG----KWG EHELDYLLFIVRD-- VKVQPNPDEVAEIKY VSREELKELVKKADA   241
 8 A.thaliana 2     AQRKLFDELGIVAED VPVDEFTPLGRMLY- --KAPSDG----KWG EHEVDYLLFIVRD-- VKLQPNPDEVAEIKY VSREELKELVKKADA   190
 9 H.pluvialis 1    AIRKLEHELGIPAHQ LPASAFRFLTRLHYC AADVQPAATQSALWG EHEMDYILFIRAN-- VTLAPNPDEVDEVRY VTQEELRQMMQP---   247
10 H.pluvialis 2    AIRKLEHELGIPAHQ LPASAFRFLTRLHYC AADVQPAATQSALWG EHEMDYILFIRAN-- VTLAPNPDEVDEVRY VTQEELRQMMQP---   259
11 C.reinhardtii 1  AVRKLQHELGIPPEQ VPASSFSFLTRLHYC AADTATHG-PAAEWG EHEVDYVLFVRPQQP VSLQPNPDEVDATRY VTLPELQSMMA----   259

                    271        285 286        300 301        315 316
 1 T.erecta 1       GEEGLKLSPWFRLVV DN--FLFKWWDHVQK GTL---TEAIDMKTI HKLI--    232  Tagetes erecta (marigold)
 2 L.sativa 1       GEEGVKLSPWFKLIV DN--FLFQWWDRLHK GTL---TEAIDMKTI HKLT--    280  Lactuca sativa (romaine lettuce)
 3 L.sativa 2       GEEGLKLSPWFRLVV DN--FLFKWWDHVQK GTL---NEAIDMKTI H-----    229  Lactuca sativa (romaine lettuce)
 4 A.palaestina 2   GEEGIKLSPWFRLVV DN--FLFKWWDHVEE GKI---KDVADMKTI HKLT--    295  Adonis palaestina (pheasant's eye)
 5 A.palaestina 1   GEEGLKLSPWFRLVV DN--FLFKWWDHVEQ GTI---KEVADMKTI HKLT--    234  Adonis palaestina (pheasant's eye)
 6 O.sativa 1       GEEGLKLSPWFRLVV DN--FLMGWWDHVEK GTL---NEAVDMETI HKLK--    238  Oryza sativa (rice)
 7 A.thaliana 1     GEEGLKLSPWFRLVV DN--FLMKWWDHVEK GTL---VEAIDMKTI HKL---    284  Arabidopsis thaliana
 8 A.thaliana 2     GDEAVKLSPWFRLVV DN--FLMKWWDHVEK GTI---TEAADMKTI HKL---    233  Arabidopsis thaliana
 9 H.pluvialis 1    -DNGLQWSPWFRIIA AR--FLERWWADLDA ALN--TDKHEDWGTV HHINEA    293  Haematococcus pluvialis
10 H.pluvialis 2    -DNGLQWSPWFRIIA AR--FLERWWADLDA ALN--TDKHEDWGTV HHINEA    305  Haematococcus pluvialis
11 C.reinhardtii 1  -DPGLSWSPWFRILA TQPAFLPAWWGDLKR RWRPGGSRLSDWGTI HRVM--    307  Chlamydomonas reinhardtii
```

FIG. 22B

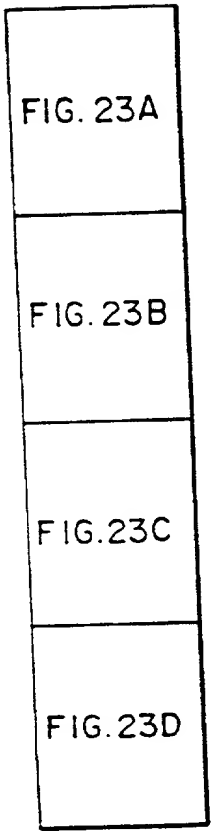


FIG. 23

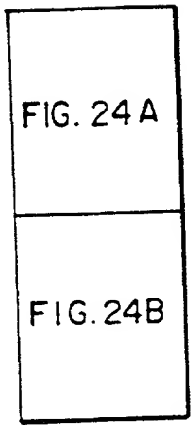


FIG. 24

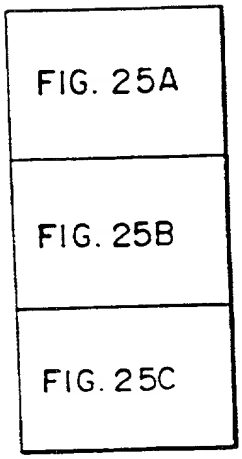


FIG. 25

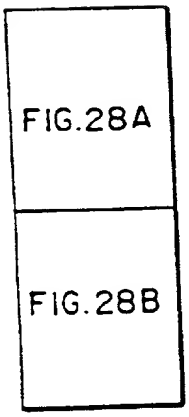


FIG. 28

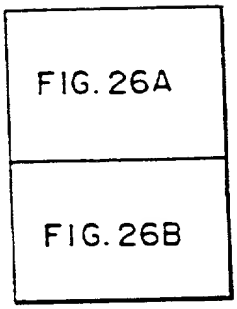


FIG. 26

FIG. 23A

Comparison using GAP program of the Genetics Computer Group

Gap Weight:	50	Average match:	10.000
Length Weight:	3	Average Mismatch:	0.000
Quality:	17392	Length:	1904
Ratio:	9.411	Gaps:	3
Percent Similarity:	95.331	Percent Identity:	95.331

Match display thresholds for the alignment(s):
| = IDENTITY : = 5 . = 1

Adonis palaestina ε-cyclase #3 x *Adonis palaestina* ε-cyclase #5

```
   1 gagagaaaaagagtgttatattaatgttactgtcgcattcttgcaacac. 49
     |||  |||||||  |||||||||||||||||||||||||||||||||||
   1 ......aaaggagtgttctattaatgttactgtcgcattcttgcaacact 44

  50 .atattcagactccattttcttgttttctcttcaaaacaacaaactaatg 98
     |||||||  |||||||||||  |||||||||||||||||||||||||||
  45 tatattcaaactccattttcttcttttctcttcaaaacaacaaactaatg 94

  99 tga.cggagtatctagctatggaactacttggtgttcgcaacctcatctc 147
     |||  |  ||||||||  |||||||||||||||||||||||||||||||
  95 tgagcagagtatctggctatggaactacttggtgttcgcaacctcatctc 144

 148 ttcttgccctgtctggacttttggaacaagaaaccttagtagttcaaaac 197
     ||||||||||||  |||||||||||||||||||||||||||||||||||
 145 ttcttgccctgtgtggacttttggaacaagaaaccttagtagttcaaaac 194

 198 tagcttataacatacatcgatatggttcttcttgtagagtagattttcaa 247
     |||||||||||||||||||||||||||||||||||||||||||||||||
 195 tagcttataacatacatcgatatggttcttcttgtagagtagattttcaa 244

 248 gtgagggctgatggtggaagcgggagtagaacttctgttgcttataaaga 297
     |||||  |||||||||||||||||||||||  |||||||||||||||||
 245 gtgagagctgatggtggaagcgggagtagaagttctgttgcttataaaga 294

 298 gggttttgtggacgaggaggattttatcaaagctggtggttctgagcttt 347
     ||||||||||||  |||||||||||||||||||||||||||||||||||
 295 gggttttgtggatgaagaggattttatcaaagctggtggttctgagcttt 344

 348 tgtttgtccaaatgcagcaaacaaagtctatggagaaacaggccaagctc 397
     |||||||||||||||||||||||||||||||||||||||||||||||||
 345 tgtttgtccaaatgcagcaaacaaagtctatggagaaacaggccaagctc 394
```

FIG. 23B

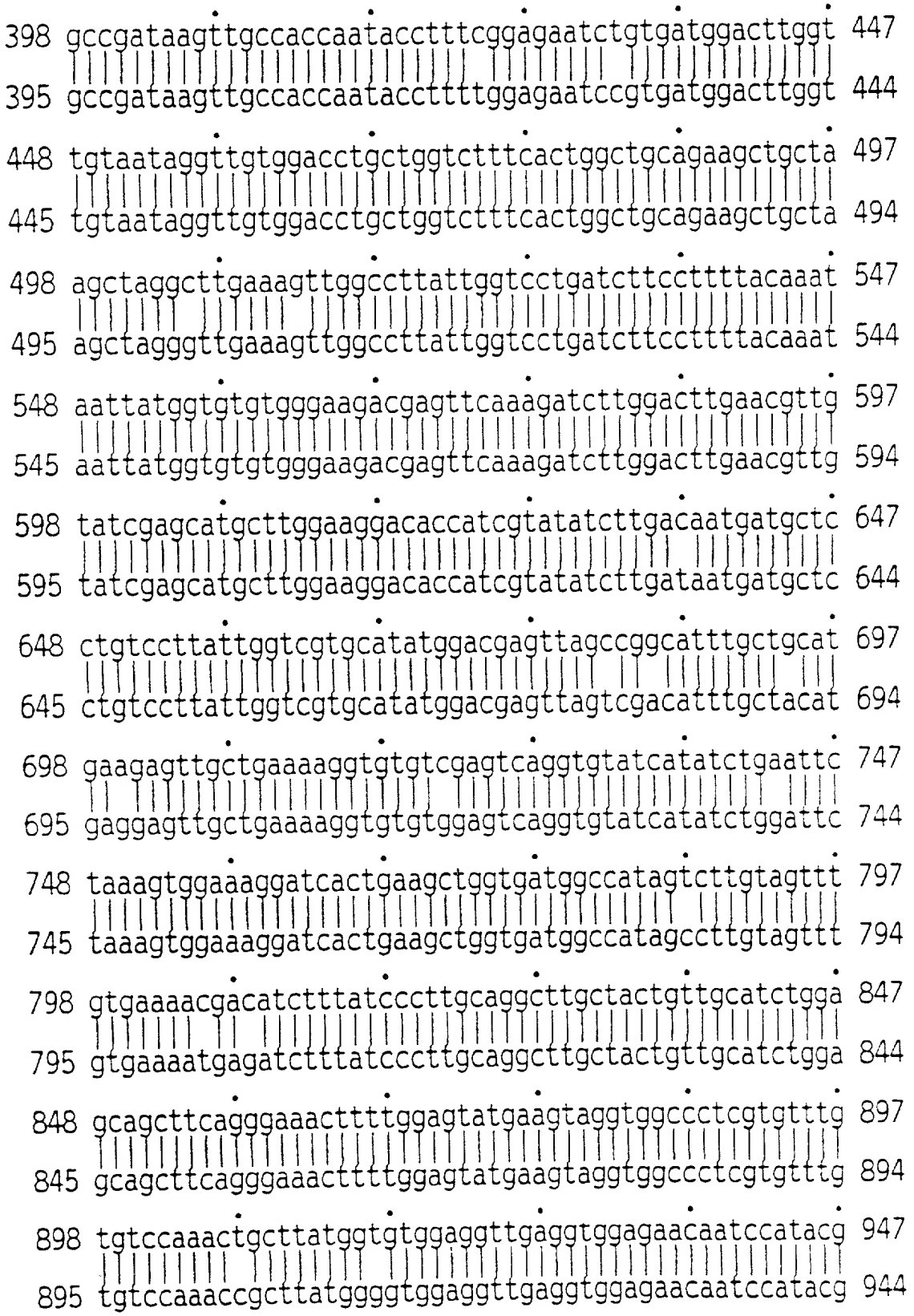


FIG. 23C

```
948 atcccaacttaatggtatttatggactacagagactatatgcaacagaaa 997
    ||||||||||||||||||| ||||||||||||||||||||||||||||||
945 atcccaacttaatggtattcatggactacagagactatatgcaacagaaa 994

998 ttacagtgctcggaagaagaatatccaacatttctctatgtcatgcccat 1047
    ||||||||||||||||||||||||||||||||||||||||||||||||||
995 ttacagtgctcggaagaagaatatccaacatttctctatgtcatgcccat 1044

1048 gtcgccaacaagactttttttgaggaaacctgtttggcctcaaaagatg 1097
     ||||||||||||||||||||||||||||||||||||||||||||||||
1045 gtcgccaacaagactttttttgaggaaacctgtttggcctcaaaagatg 1094

1098 ccatgcctttcgatctactgaagagaaaactaatgtcacgattgaagact 1147
     ||||||| ||||||||||||||||||||| ||||||||||||||||||
1095 ccatgccattcgatctactgaagagaaaactgatgtcacgattgaagact 1144

1148 ctgggtatccaagttacaaaaatttatgaagaggaatggtcttatattcc 1197
     |||||||||||||||||||| |||||||||||||||| ||||||||
1145 ctgggtatccaagttacaaaagtttatgaagaggaatggtcatatattcc 1194

1198 tgttgggggttctttaccaaacacagagcaaaagaacctagcatttggtg 1247
     |||||| ||||||||||||||||||||||||||||||||||||||||
1195 tgttggtggttctttaccaaacacagagcaaaagaacctagcatttggtg 1244

1248 ctgcagcaagcatggtgcatccagcaacaggctattcggttgtacgatca 1297
     ||||||||||||||||||||||||||||||||||||||||||| |||
1245 ctgcagcaagcatggtgcatccagcaacaggctattcggttgtacggtca 1294

1298 ctatcagaagctccaaaatatgcttctgtaattgcaaagattttgaagca 1347
     || ||||||||||||||||||||||||||||||||||||||||||||
1295 ctgtcagaagctccaaaatatgcttctgtaattgcaaagattttgaagca 1344

1348 agataactctgcatatgtggtttctggacaaagcagtgcagtaaacattt 1397
     |||||||||||| |||||||||| |||||||| ||||||||||||||
1345 agataactctgcgtatgtggtttctggacaaagtagtgcagtaaacattt 1394

1398 caatgcaagcatggagcagtctttggccaaaggagcgaaaacgtcaaaga 1447
     ||||||||||||||||||||||||||||||||||||||||||||||||
1395 caatgcaagcatggagcagtctttggccaaaggagcgaaaacgtcaaaga 1444

1448 gcattctttcttttcgggttagagcttattgtgcagctagatattgaagc 1497
     |||||||||||||| || ||||||||||||||||||||||||||||||
1445 gcattctttcttttggattagagcttattgtgcagctagatattgaagc 1494
```

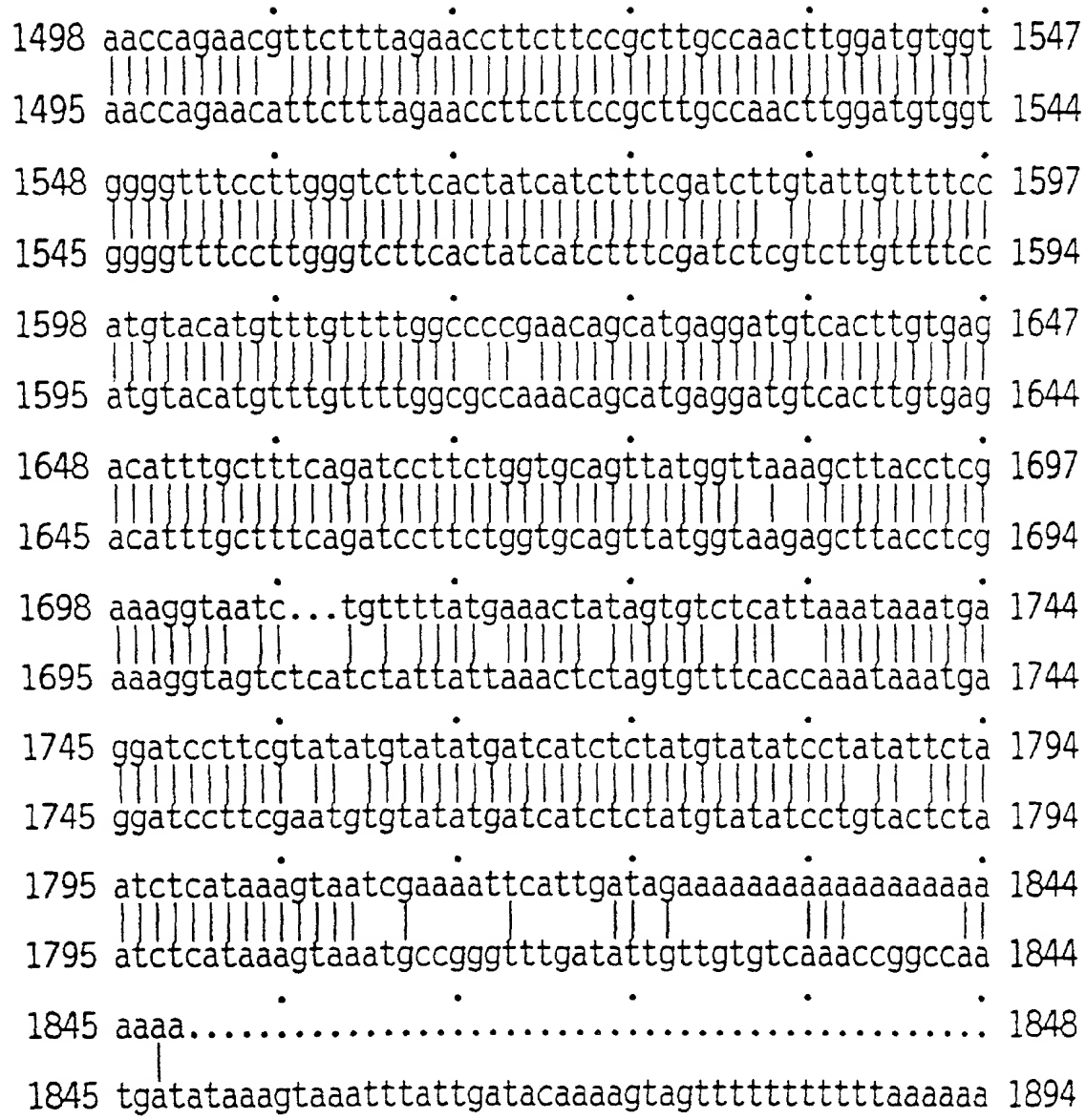

1498 aaccagaacgttctttagaaccttcttccgcttgccaacttggatgtggt 1547
1495 aaccagaacattctttagaaccttcttccgcttgccaacttggatgtggt 1544
1548 ggggtttccttgggtcttcactatcatctttcgatcttgtattgttttcc 1597
1545 ggggtttccttgggtcttcactatcatctttcgatctcgtcttgttttcc 1594
1598 atgtacatgtttgttttggccccgaacagcatgaggatgtcacttgtgag 1647
1595 atgtacatgtttgttttggcgccaaacagcatgaggatgtcacttgtgag 1644
1648 acatttgctttcagatccttctggtgcagttatggttaaagcttacctcg 1697
1645 acatttgctttcagatccttctggtgcagttatggtaagagcttacctcg 1694
1698 aaaggtaatc...tgttttatgaaactatagtgtctcattaaataaatga 1744
1695 aaaggtagtctcatctattattaaactctagtgtttcaccaaataaatga 1744
1745 ggatccttcgtatatgtatatgatcatctctatgtatatcctatattcta 1794
1745 ggatccttcgaatgtgtatatgatcatctctatgtatatcctgtactcta 1794
1795 atctcataaagtaatcgaaaattcattgatagaaaaaaaaaaaaaaaaaa 1844
1795 atctcataaagtaaatgccgggtttgatattgttgtgtcaaaccggccaa 1844
1845 aaaa..... 1848
1845 tgatataaagtaaatttattgatacaaaagtagttttttttttaaaaaa 1894

FIG. 23D

FIG. 24A

GAP program of Genetics Computer Group
blosum62.cmp

Gap Weight:	12	Average Match:	2.912
Length Weight:	4	Average Mismatch:	-2.003
Quality:	2728	Length:	530
Ratio:	5,147	Gaps:	0
Percent Similarity:	99,623	Percent Identity:	99.057

Match display thresholds for the alignment(s):
| = IDENTITY : = 2 . = 1

Adonis palaestina ε-cyclase #3 x *Adonis palaestina* ε-cyclase #5

```
   1 MELLGVRNLISSCPVWTFGTRNLSSSKLAYNIHRYGSSCRVDFQVRADGG 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
   1 MELLGVRNLISSCPVWTFGTRNLSSSKLAYNIHRYGSSCRVDFQVRADGG 50

  51 SGSRTSVAYKEGFVDEEDFIKAGGSELLFVQMQQTKSMEKQAKLADKLPP 100
     ||||.|||||||||||||||||||||||||||||||||||||||||||||
  51 SGSRSSVAYKEGFVDEEDFIKAGGSELLFVQMQQTKSMEKQAKLADKLPP 100

 101 IPFGESVMDLVVIGCGPAGLSLAAEAAKLGLKVGLIGPDLPFTNNYGVWE 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 101 IPFGESVMDLVVIGCGPAGLSLAAEAAKLGLKVGLIGPDLPFTNNYGVWE 150

 151 DEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKR 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 151 DEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKR 200

 201 CVESGVSYLNSKVERITEAGDGHSLVVCENDIFIPCRLATVASGAASGKL 250
     |||||||||.|||||||||||||||||:||||||||||||||||||||||
 201 CVESGVSYLDSKVERITEAGDGHSLVVCENEIFIPCRLATVASGAASGKL 250

 251 LEYEVGGPRVCVQTAYGVEVEVENNPYDPNLMVFMDYRDYMQQKLQCSEE 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 251 LEYEVGGPRVCVQTAYGVEVEVENNPYDPNLMVFMDYRDYMQQKLQCSEE 300

 301 EYPTFLYVMPMSPTRLFFEETCLASKDAMPFDLLKRKLMSRLKTLGIQVT 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 301 EYPTFLYVMPMSPTRLFFEETCLASKDAMPFDLLKRKLMSRLKTLGIQVT 350
```

```
                .         .         .         .         .
351 KIYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAPK 400
    |:||||||||||||||||||||||||||||||||||||||||||||||||||
351 KVYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAPK 400

                .         .         .         .         .
401 YASVIAKILKQDNSAYVVSGQSSAVNISMQAWSSLWPKERKRQRAFFLFG 450
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
401 YASVIAKILKQDNSAYVVSGQSSAVNISMQAWSSLWPKERKRQRAFFLFG 450

                .         .         .         .         .
451 LELIVQLDIEATRTFFRTFFRLPTWMWWGFLGSSLSSFDLVLFSMYMFVL 500
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
451 LELIVQLDIEATRTFFRTFFRLPTWMWWGFLGSSLSSFDLVLFSMYMFVL 500

                .         .         .
501 APNSMRMSLVRHLLSDPSGAVMVKAYLER* 530
    ||||||||||||||||||||||:|||||||
501 APNSMRMSLVRHLLSDPSGAVMVRAYLER* 530
```

FIG. 24B

FIG. 25A

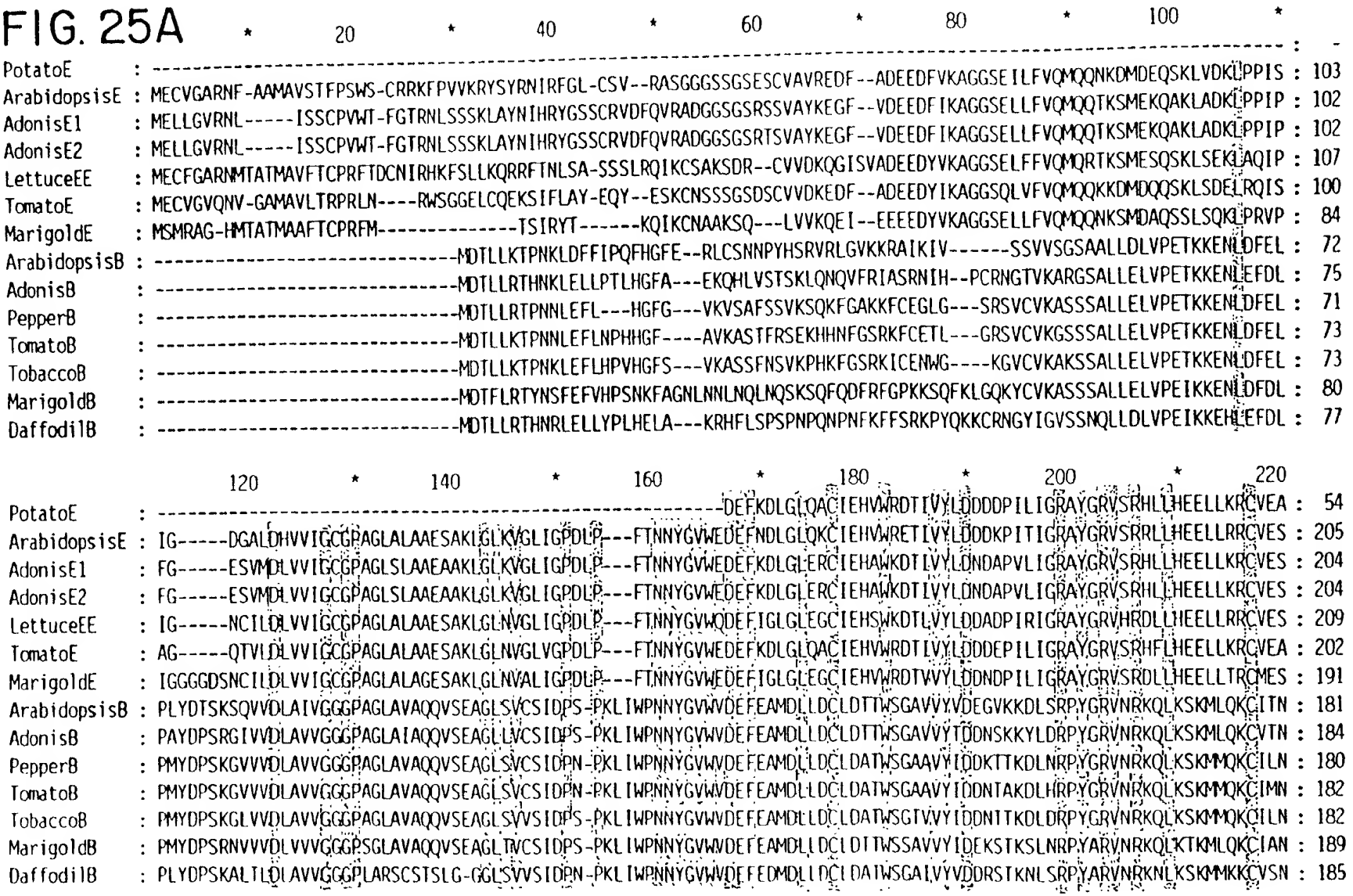


FIG. 25B

FIG. 25C

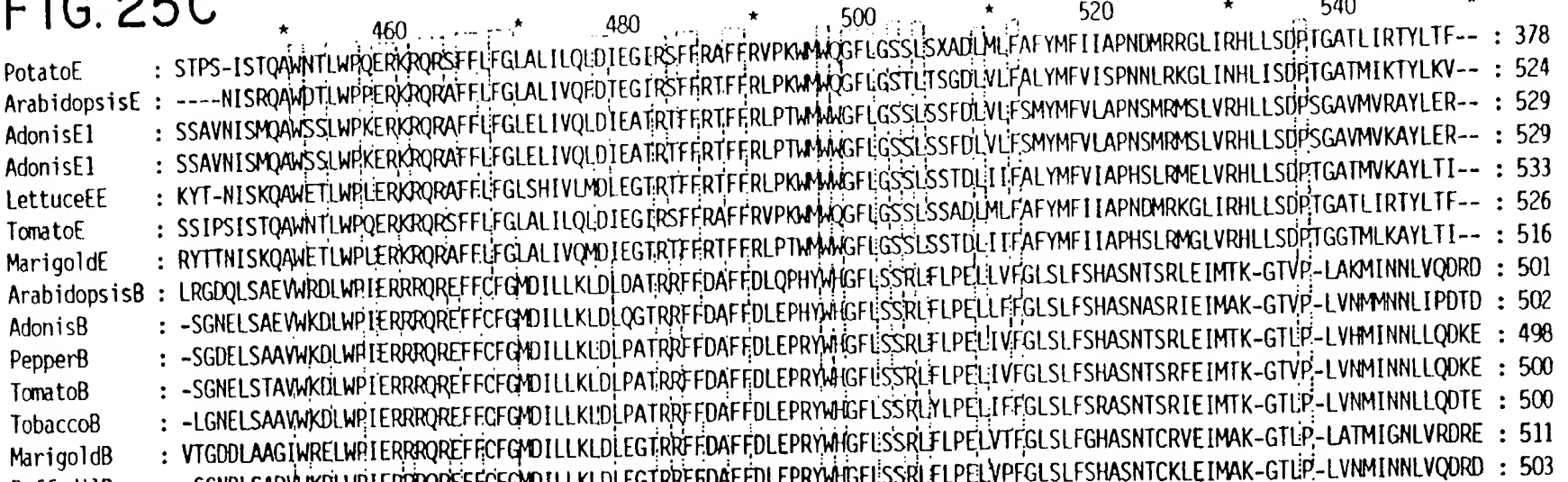
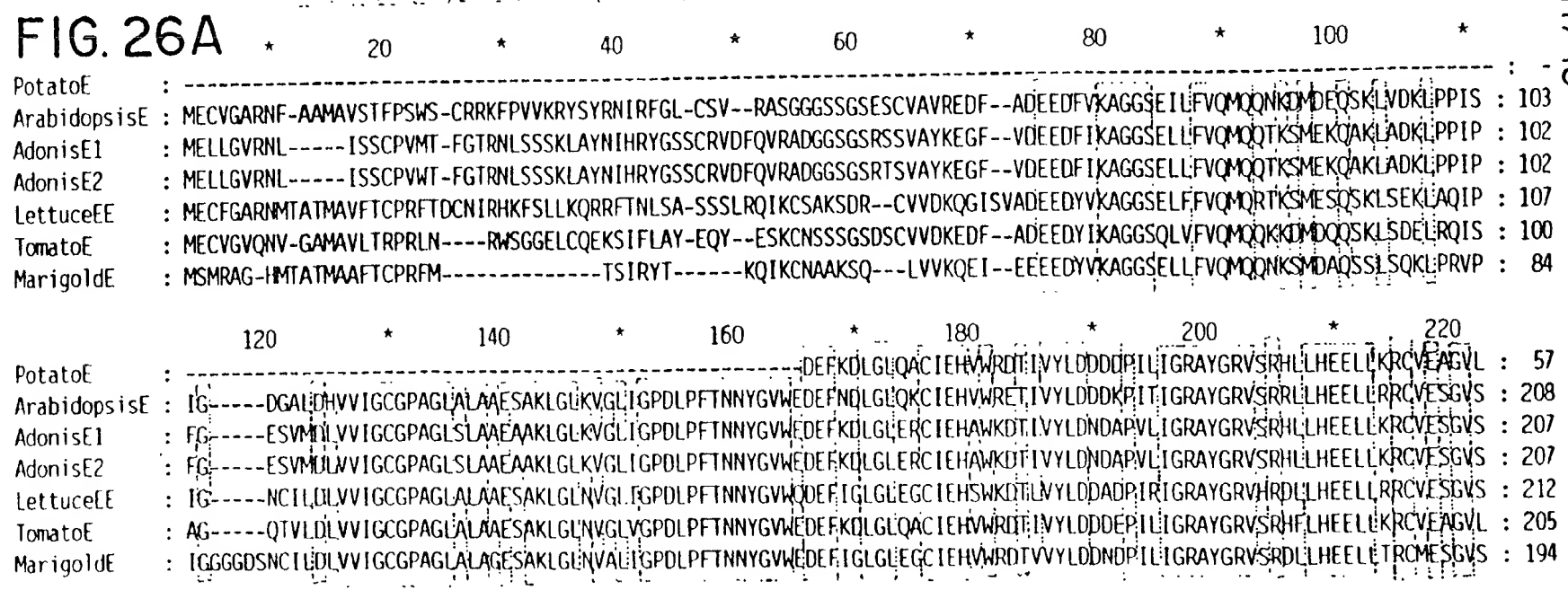


FIG. 26A



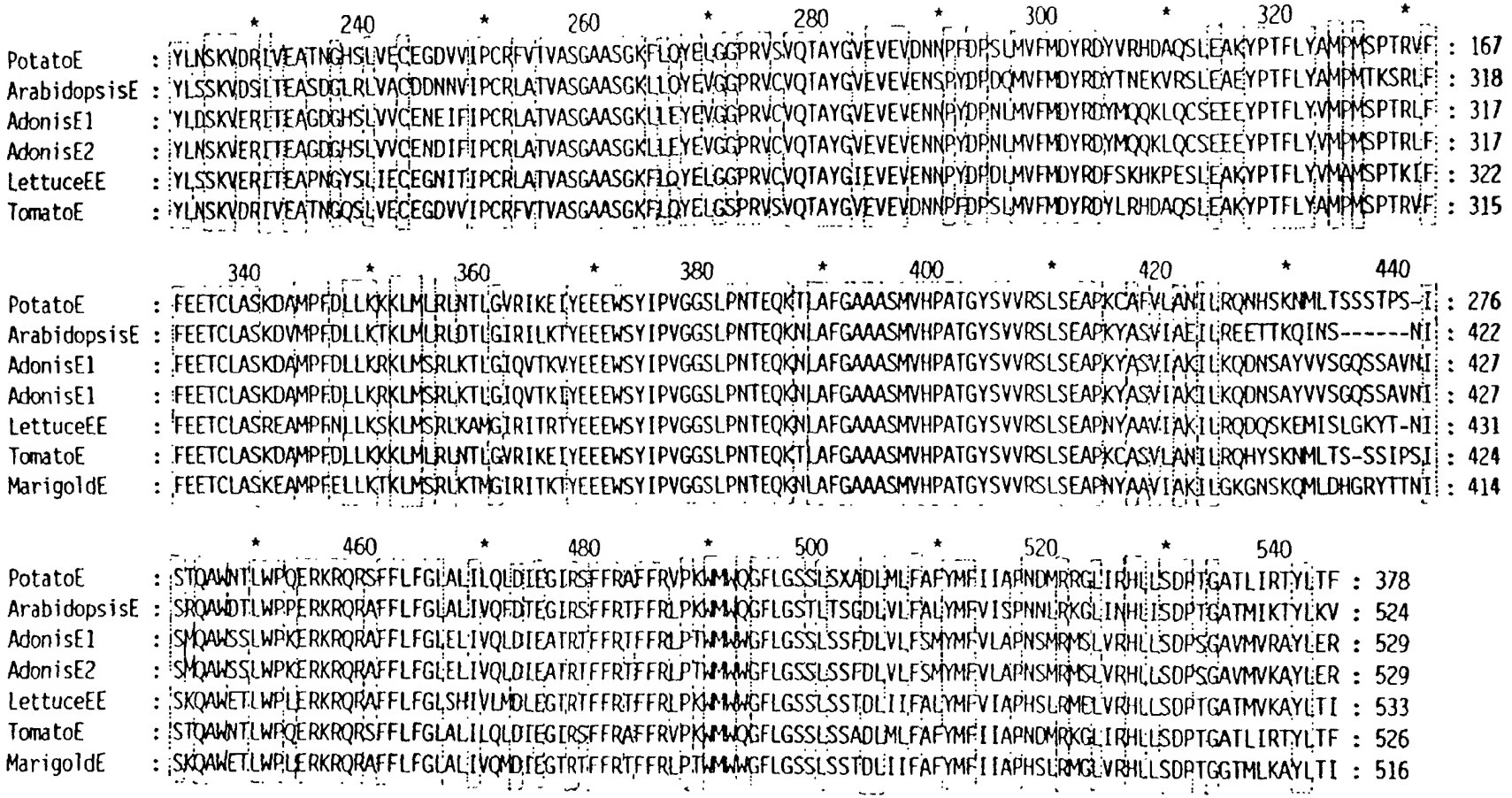


FIG. 26B

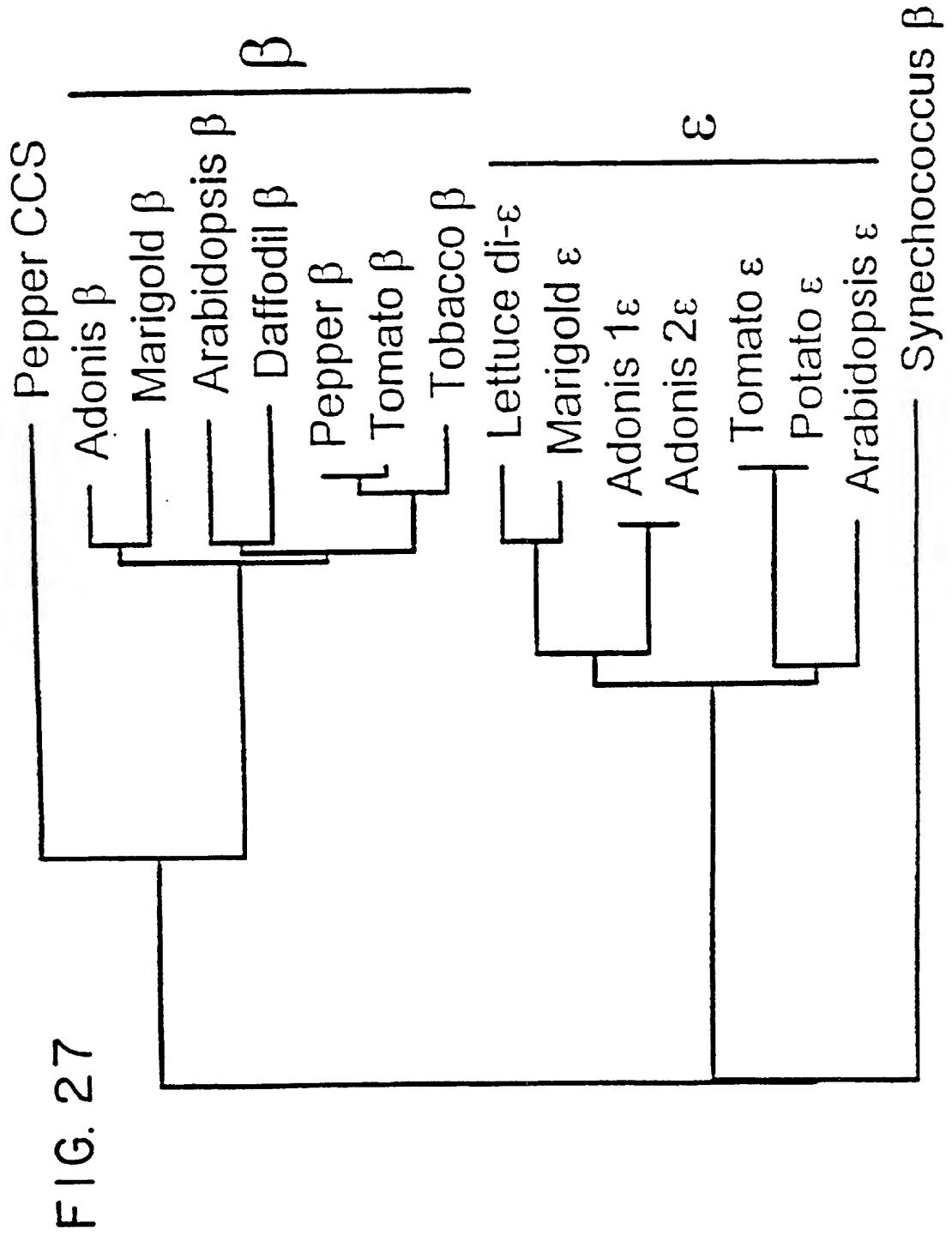

β
ε
Pepper CCS
Adonis β
Marigold β
Arabidopsis β
Daffodil β
Pepper β
Tomato β
Tobacco β
Lettuce di-ε
Marigold ε
Adonis 1ε
Adonis 2ε
Tomato ε
Potato ε
Arabidopsis ε
Synechococcus β

FIG. 27

FIG. 28A

GAP of: Arabidopsis epsilon cyclase to Lettuce epsilon cyclase

Gap Weight:	12	Average Match:	2.912
Length Weight:	4	Average Mismatch:	-2.003
Quality:	1837	Length:	534
Ratio:	3.499	Gaps:	3
Percent Similarity:	76.381	Percent Identity:	69.905

Match display thresholds for the alignment(s):
| = IDENTITY : = 2 . = 1

Arabidopsis x Lettuce

```
  1 MECVGARNF.AAMAVSTFPSW...SCRRKFPVVKRYSYRNIRFGLCSVRA 46
    ||| |||| | ||| | | .  . | || ..|.  : |:        .
  1 MECFGARNMTATMAVFTCPRFTDCNIRHKFSLLKQRRFTNLSASSSLRQI 50

 47 SGGGSSGSESCVAVREDFADEEDFVKAGGSEILFVQMQQNKDMDEQSKLV 96
       |             |||||:|||||||: |||||. | |: ||||
 51 KCSAKSDRCVVDKQGISVADEEDYVKAGGSELFFVQMQRTKSMESQSKLS 100

 97 DKLPPISIGDGALDHVVIGCGPAGLALAAESAKLGLKVGLIGPDLPFTNN 146
    :||  | ||.  || |||||||||||||||||||| ||||||||||||
101 EKLAQIPIGNCILDLVVIGCGPAGLALAAESAKLGLNVGLIGPDLPFTNN 150

147 YGVWEDEFNDLGLQKCIEHVWRETIVYLDDDKPITIGRAYGRVSRRLLHE 196
    ||||:|||  |||: |||| |::|:|||||  || |||||||| | ||||
151 YGVWQDEFIGLGLEGCIEHSWKDTLVYLDDADPIRIGRAYGRVHRDLLHE 200

197 ELLRRCVESGVSYLSSKVDSITEASDGLRLVACDDNNAIPCRLATVASGA 246
    ||||||||||||||||||: |||| .|  |: |: |  ||||||||||||
201 ELLRRCVESGVSYLSSKVERITEAPNGYSLIECEGNITIPCRLATVASGA 250

247 ASGKLLQYEVGGPRVCVQTAYGVEVEVENSPYDPDQMVFMDYRDYTNEKV 296
    ||||  |:||.||||||||||||:||||||.|||||  |||||||:.  |
251 ASGKFLEYELGGPRVCVQTAYGIEVEVENNPYDPDLMVFMDYRDFSKHKP 300

297 RSLEAEYPTFLYAMPMTKSRLFFEETCLASKDVMPFDLLKTKLMLRLDTL 346
     ||||.|||||| | |. .::|||||||||:: |||.|||.|||  ||  :
301 ESLEAKYPTFLYVMAMSPTKIFFEETCLASREAMPFNLLKSKLMSRLKAM 350
```

FIG. 28B

```
347 GIRILKTYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSL 396
    ||||  :||||||||||||||||||||||||||||||||||||||||||||
351 GIRITRTYEEEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSL 400

397 SEAPKYASVIAEILREETTKQINS.....NISRQAWDTLWPPERKRQRAF 441
    ||||  ||.|||.|||:: .|:. |     |||:|||:|||| |||||||||
401 SEAPNYAAVIAKILRQDQSKEMISLGKYTNISKQAWETLWPLERKRQRAF 450

442 FLFGLALIVQFDTEGIRSFFRTFFRLPKWMWQGFLGSTLTSGDLVLFALY 491
    |||||. ||  | || |.|||||||||||||| |||||.|.| ||::||||
451 FLFGLSHIVLXDLEGTRTFFRTFFRLPKWMWWGFLGSSLSSTDLIIFALY 500

492 MFVISPNNLRKGLINHLISDPTGATMIKTYLKV* 525
    ||||.|..||  |: ||:|||||||:| || :|
501 MFVIAPHSLRMELVRHLLSDPTGATMVKAYLTI* 534
```

Docket No. 108172-00022

ARENT FOX KINTNER PLOTKIN & KAHN, PLLC
Nikaido, Marmelstein, Murray & Oram Intellectual Property Group

Declaration For U.S. Patent Application

As a below named inventor, I hereby declare that:
My residence, post office address and citizenship are as stated below my name.
I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention **entitled**
(Insert Title) GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF

the specification of which is attached hereto unless the following box is checked:

☒ was filed on June 2, 1999 As PCT International Application
Number PCT/US99/12121 and was amended on _____
and/or was filed on December 4, 2000 As U.S. Patent Application
Number _____ and was amended on _____ .

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claim(s), as amended by any amendment referred to above.
I acknowledge the duty to disclose information which is material to patentability as defined in 37 C.F.R. §1.56.
I hereby claim foreign priority benefits under 35 U.S.C. §119(a)-(d) or §365(b) of any foreign application(s) for patent or inventor's certificate, or §365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below any foreign application for patent or inventor's certificate or PCT International Application having a filing date before that of the application(s) for which priority is claimed:

(List prior foreign applications)	(Number)	(Country)	(Day/Month/Year Filed)	Priority Claimed
				☐ Yes ☐ No
				☐ Yes ☐ No
				☐ Yes ☐ No

I hereby claim the benefit under 35 U.S.C. §119(e) of any United States provisional application(s) listed below.

(Application Number)	(Filing Date)

☐ See attached list for additional prior foreign or provisional applications.

I hereby claim the benefit under 35 U.S.C. §120 of any United States application(s) or §365(c) of any PCT International application(s) designating the United States of America listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior application(s) (U.S. or PCT) in the manner provided by the first paragraph of 35, U.S.C. §112, I acknowledge the duty to disclose information which is material to patentability as defined in 37 C.F.R. §1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application.

(List prior U.S. Applications or PCT International applications designating the U S)	(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)
	09/088,724	June 2, 1998	
	09/088,725	June 2, 1998	

And I hereby appoint the firm of Arent Fox, Customer Number 004372 including as principal attorneys: Robert B. Murray, Reg. No. 22,980; Charles M. Marmelstein, Reg. No. 25,895; George E. Oram, Jr., Reg. No. 27,931; Douglas H. Goldhush, Reg. No. 33,125; David T. Nikaido, Reg. No. 22,663; Richard J. Berman, Reg. No. 39,107; King L. Wong, Reg. No. 37,500; James A. Poulos, III, Reg. No. 31,714; Murat Ozgu, Reg. No. 44,275; Robert K. Carpenter, Reg. No. 34,794; Gregory B. Kang, Reg. No. 45,273; Rustan Hill, Reg. No. 37,351; Carl Schaukowitch, Reg. No. 29,211; Kevin Turner, Reg. No. 43,437; Rhonda C. Barton, Reg. No. P47,271 and Hans J. Crosby, Reg. No. 44,634.

Please direct all communications to the following address:
ARENT FOX KINTNER PLOTKIN & KAHN, PLLC
1050 Connecticut Avenue, N.W., Suite 600
Washington, D.C. 20036-5339
Telephone No. (202) 857-6000; Facsimile No. (202) 638-4810

The undersigned hereby authorizes the U.S. attorneys named herein to accept and follow instructions from the undersigned's assignee, if any, and/or, if the undersigned is not a resident of the United States, the undersigned's domestic attorney, patent attorney or patent agent, as to any action to be take in the Patent and Trademark Office regarding this application without direct communication between the U.S. attorneys and the undersigned. In the event of a change in the person(s) from whom instructions may be taken, the U.S. attorneys named herein will be so notified by the undersigned

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first inventor Francis X. CUNNINGHAM, Jr.
Inventor's signature ______________________ ______________
Date
Residence 2727 Washington Avenue, Chevy Chase, MD 20815
Citizenship American
Post Office Address Same as the above

2-00
Full name of sole or second inventor Zairen SUN
Inventor's signature Zairen Sun 1/10/2001
Date
Residence 3~~405 Tulane Drive, #22, Hyattsville, MD 20783~~ 1083 Copperstone Ct. Rockville, MD 20852
Citizenship permant resident, Chinese, P.R. China
Post Office Address Same as the above

Full name of sole or third inventor ______________________
Inventor's signature ______________________ ______________
Date
Residence ______________________
Citizenship ______________________
Post Office Address ______________________

Docket No. 108172-00022

ARENT FOX KINTNER PLOTKIN & KAHN, PLLC
Nikaido, Marmelstein, Murray & Oram Intellectual Property Group

Declaration For U.S. Patent Application

As a below named inventor, I hereby declare that:
My residence, post office address and citizenship are as stated below my name.
I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention **entitled**
(Insert Title) GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND METHODS OF USE THEREOF

the specification of which is attached hereto unless the following box is checked:

☒ was filed on June 2, 1999 As PCT International Application
Number PCT/US99/12121 and was amended on _____
and/or was filed on December 4, 2000 As U.S. Patent Application
Number _____ and was amended on _____

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claim(s), as amended by any amendment referred to above.
I acknowledge the duty to disclose information which is material to patentability as defined in 37 C.F.R. §1.56.
I hereby claim foreign priority benefits under 35 U.S.C. §119(a)-(d) or §365(b) of any foreign application(s) for patent or inventor's certificate, or §365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below any foreign application for patent or inventor's certificate or PCT International Application having a filing date before that of the application(s) for which priority is claimed:

(List prior foreign applications)	(Number)	(Country)	(Day/Month/Year Filed)	Priority Claimed
				☐ Yes ☐ No
				☐ Yes ☐ No
				☐ Yes ☐ No

I hereby claim the benefit under 35 U.S.C. §119(e) of any United States provisional application(s) listed below.

(Application Number)	(Filing Date)

☐ See attached list for additional prior foreign or provisional applications.

I hereby claim the benefit under 35 U.S.C. §120 of any United States application(s) or §365(c) of any PCT International application(s) designating the United States of America listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior application(s) (U.S. or PCT) in the manner provided by the first paragraph of 35, U.S.C. §112, I acknowledge the duty to disclose information which is material to patentability as defined in 37 C.F.R. §1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application.

(List prior U.S. Applications or PCT International applications designating the U S)	(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)
	09/088,724	June 2, 1998	
	09/088,725	June 2, 1998	

And I hereby appoint the firm of Arent Fox, Customer Number 004372 including as principal attorneys: Robert B. Murray, Reg. No. 22,980; Charles M. Marmelstein, Reg. No. 25,895; George E. Oram, Jr., Reg. No. 27,931; Douglas H. Goldhush, Reg. ~~No. 33,125~~; David T. Nikaido, Reg. ~~No. 22,663~~; Richard J. Berman, Reg. No. 39,107; King L. Wong, Reg. No. 37,500; James A. Poulos, III, Reg. No. 31,714; Murat Ozgu, Reg. ~~No. 44,275~~; Robert K. Carpenter, Reg. No. 34,794; Gregory B. Kang, Reg. ~~No. 45,273~~; Rustan Hill, Reg. ~~No. 37,351~~; Carl Schaukowitch, Reg. ~~No. 29,211~~; Kevin Turner, Reg. No. 43,437; Rhonda C. Barton, Reg. ~~No. P47,271~~ and Hans J. Crosby, Reg. No. 44,634.

Please direct all communications to the following address:

ARENT FOX KINTNER PLOTKIN & KAHN, PLLC
1050 Connecticut Avenue, N.W., Suite 600
Washington, D.C. 20036-5339
Telephone No. (202) 857-6000; Facsimile No. (202) 638-4810

The undersigned hereby authorizes the U.S. attorneys named herein to accept and follow instructions from the undersigned's assignee, if any, and/or, if the undersigned is not a resident of the United States, the undersigned's domestic attorney, patent attorney or patent agent, as to any action to be take in the Patent and Trademark Office regarding this application without direct communication between the U.S. attorneys and the undersigned. In the event of a change in the person(s) from whom instructions may be taken, the U.S. attorneys named herein will be so notified by the undersigned

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

1-DD

Full name of sole or first inventor Francis X. CUNNINGHAM, Jr.
Inventor's signature Francis X Cunningham, Jr. Feb. 16, 2001
Date

Residence 2727 Washington Avenue, Chevy Chase, MD 20815
Citizenship American
Post Office Address Same as the above

Full name of sole or second inventor Zairen SUN
Inventor's signature _____ _____
Date

Residence 3405 Tulane Drive, #22, Hyattsville, MD 20783
Citizenship _____
Post Office Address Same as the above

Full name of sole or third inventor _____
Inventor's signature _____ _____
Date

Residence _____
Citizenship _____
Post Office Address _____